

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2004, 04:08:52 ; Search time 46 Seconds
(without alignments)
3709.817 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494

Sequence: 1 MTKFKKILIVFAVIALSSG.....KGLMLTLITNGKQQLVLP 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1049977 seqs, 258955339 residues

Total number of hits satisfying chosen parameters: 1049977

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	7.0	896	12 US-10-282-122A-45584	Sequence 45584, A
2	130.5	3.7	681	12 US-10-282-122A-71561	Sequence 71561, A
3	128	3.7	25	9 US-09-802-285-3	Sequence 3, Appli
4	128	3.7	25	14 US-10-291-337-3	Sequence 3, Appli
5	126.5	3.6	696	12 US-10-282-122A-70746	Sequence 70746, A
6	123	3.5	532	15 US-10-369-493-22074	Sequence 22074, A
7	120	3.4	677	15 US-10-369-493-23212	Sequence 23212, A
8	118.5	3.4	579	15 US-10-369-493-23115	Sequence 23115, Ap
9	118.5	3.4	579	15 US-10-369-493-2321	Sequence 2321, Ap
10	117.5	3.4	2902	12 US-10-282-122A-59042	Sequence 59042, A
11	117	3.3	1163	15 US-10-452-024-107	Sequence 107, App
12	117	3.3	1250	10 US-09-769-736-10	Sequence 10, Appl
13	116	3.3	919	12 US-10-282-122A-70737	Sequence 70737, A
14	116	3.3	4563	10 US-09-802-640-32	Sequence 32, Appl
15	115.5	3.3	987	12 US-10-282-122A-48445	Sequence 48445, A

Sequence 1532, Ap
Sequence 370, Appli
Sequence 8, Appli
Sequence 388, App
Sequence 13471, A
Sequence 5864, Ap
Sequence 46565, A
Sequence 6, Appli
Sequence 3481, Ap
Sequence 3302, Ap
Sequence 77946, A
Sequence 10942, A
Sequence 5665, Ap
Sequence 12206, A
Sequence 128, App
Sequence 128, App
Sequence 72, Appl
Sequence 72, Appl
Sequence 47271, A
Sequence 52070, A
Sequence 22, Appl
Sequence 654, App
Sequence 128, App
Sequence 92, Appl
Sequence 204, App
Sequence 12, Appl
Sequence 47599, A
Sequence 43071, A
Sequence 7128, Ap
Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-10-282-122A-45584
; Sequence 45584, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636

```
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 45584
; LENGTH: 896
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-10-282-122A-45584

Query Match
Best Local Similarity 7.0%; Score 243; DB 12; Length 896;
Matches 132; Conservative 93; Mismatches 250; Indels 120; Gaps 29;

QY 113 XGYFDYKIDNQWPKVDNEVRWQ-----LHRVKKWQAMALVHATGDEKYAREWVYQY 167
Db 47 FGSSSYNLIHW-----NEQREGYRLIHGHTFLGCLIAAYNDTGDMYIKKSIELI 99
QY 168 SDWARKNPLGLSQDNKDFWVRPLEVSDRVQSLPPTFSLFVNSPAPTPAFME-----FLN 222
Db 100 KDWNNHSEFELHQHSMAR-----HDETTAL--RLQWLRFYFTRQVLSSEELILEK 150
QY 223 SYHQOAYLSTHY--AEQGNHRLFEAQNLFAGVSFPPEFK-DSP---RWRQTGISVLNTE 276
Db 151 SMEDTAKLLSDFFHATNTNMGFQ--DRALLTYASY--FKGENPSLEKYIKLAVTRLKDY 207
QY 277 IKQVYADGMOFELSPIYHVAAIDIFLKAYSARKVNLKEFPQSPQSYVOTVENMIMA----- 332
Db 208 PEKVFTTEGVHKEHSPSYHL-----LVASNIKKLANMKBFDRKE-VSLIFNKIYKKTTEE 260
QY 333 -LISISLPDYNTPMFGDSW--ITDKNFRMAQFASWARVFPANQAIKYPATDGKOGKAPNF 389
Db 261 YAIHIIIRDGSLPDCITEANLVGNKY-----DLYESDQVL-VYVTKGKGKAPTE 311
QY 390 LSKALSNAGFTTFRSGWDK--NATVMVLKASPPGEFHAQPDNGTFELFKGRNFTPDAGV 447
Db 312 DDKVFPKSGVAIFRNDWSKEEKATYVLFPTAAHYVDYHKHSD-----DLNL 356
QY 448 FVYSGDEALMKL--RNWYR-----QTFIHSTLIDNQNWITKARQNK-----WET 491
Db 357 YTSNGEIIIEAGNGYNYKDPFTEYAYSFAHNTLIVDGKGLPRTDQYKVKYLSDEI 416
QY 492 GNLDVLTYTNPSYENLDHORSVLFINKKYPFLVT--DRAIGBATGNLGVHMLKEDSNPVE 550
Db 417 NKDKVEATGYNLRTGVVHSEKTSYMKDEKILVVDLVKSKRHEYKLLMHWASD---IT 473
QY 551 DTKNRYVITYRDGNLMIOSL-----NADRTSLNEE--EGKVSYYNNKELKRPAPVFEKPK 605
Db 474 VHVDRDRIVELFRNNHKVMEMEVTTVTGVSIRALNEQTKPVS-----GWVFPKMG 523
QY 606 KNAQTQNFVSIVYPYDQKAPESIRENKGNDFEKK-----LNLITLTKNGKQOL 655
Db 524 EKRGK---TTFIEVDISGNVE--CITEFLKDFKLGKRDLLIPYLNLEKTFKSTRNL 573
```

RESULT 2

```
US-10-282-122A-71561
; Sequence 71561, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
```

```
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 71561
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-71561

Query Match
Best Local Similarity 3.7%; Score 130.5; DB 12; Length 681;
Matches 134; Conservative 88; Mismatches 256; Indels 193; Gaps 27;

QY 9 IIVFAVIALSGNI-LAQSSSIKDFDHNLEYSGLEK-----VNKAVAGN----- 55
Db 30 WVIFGIIVRLGYLQIAQS-----QYSQLVKNDENTVNESVPRGILDRNG 77
QY 56 ---YDDAAKALLAYYREKSKAREPDFSNAPKPAD--IRQPIDKVTFR----- 96
Db 78 KVLVDNASKLTIYTRSRKTSQKMDLTAKLSLLITMKTDKITERKQDFWIKHQDEV 137
QY 97 -----EMADKALVHQFQPHKG-----YGYFDYKGDINWQ 125
Db 138 DKLMKKESTMLNEGSITQDOYDKQLYKKVGDKQINSLSKKDLRLVALYREMSAGSTWNPQ 197
QY 126 MWPKQNEVRWQLHRVKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGLSQDNKDF 185
Db 198 --TKNEDVSEK-----EYAAVSQQLDSLPGVNTMDWDRYPYV----- 235
QY 186 VWRPLEVSDRVQSLPPTFSLFVNSPAPTPAFLEFLNSYHQADYLSHYAEQGNHRLFE 245
Db 236 -----DTLRSI---FGSVSTSSSEGI PKELTE-----QYLAKGYS----- 266
QY 246 AQNLFAVGSFPE--FKDSPRWQTGI-----SVLNTE--IKQVYADGMOFELSPIY 294
Db 267 --RNDRVKSYLEYQYEDILRGKKEMKYTTDKSGKVINSEVINPGSRGDDLQI----- 318
QY 295 HVAADIFLKAYSARKVN--LEKEFPQSVVQTVENNMALISLSLDYNTPEGDSWIT 352
Db 319 ---TIDIDLO-----KKVESLLENQIKTLRSQAKOMDNALI VQNPKNGLIILAMAGKQI 370
QY 353 DKNFRMAQF--ASWARVFPANQAIK-----YFATDGKQKAPNLFSLKALSNAFYTFR 403
Db 371 DKNGKLTIDYDLGNFTGQFAVGSVSKGTLLAGYQNNAIKVGE--EMIDEPLHFKGLTKR 428
QY 404 SGWDKATVMVLKASPPGEFHAQPDNGTFELFKGRNFTPDAGVFTVSG-----DBAI 456
Db 429 SYFNKNDKVR-----NDKEALMHSSNVYMF---KTALKLAGDPYYSGMGLPTDISAG 479
```


147 MALV-----YHATGDEKAYREWVYQYSDWARKNPLGLSQND 183
 182 LAIYREMNAGSTLDPQTIKNEVDSEKAAVQSLKPLPGVNTTMDWDEKPYG---DTL 238
 184 KFWRPPLVSDRVQSLPTFTSLFVNSPAPFTAFMEFLNSVHQADYLSHYABQGNHRL 243
 239 RGIFG--DVSTSTGIPKELT-----EQYLSKGYSRNDRVGKSYLEYQYEDV 283
 244 FEARNLPAFVSPFEKDSRWRQIGISVLNTEIKQVYADGMQFELSPIYHVAIDIFL 303
 284 LKGTQKQ-----KYTTDKSGRVISSEVLNPGSR-----GHDQLQLT-----IDL 324
 304 KAYGSAKRVN--LEKEFPQSYQVTVENMIMALISISLPDYNTPMFGSDSWITDKNFRMAQF 361
 325 Q-----KKVESLLEKQISKLRSQAKMDNALMVVNQPNKGDILAIAGKQIDKQKLUKY 379
 362 --ASWARVFPANQAIK-YFATDGKQKAPN-----FLSKALSNAGFYPRSGWDKNA----- 410
 380 DIGNFTAQYTVGSSVKGGTLLAGYQNKAINVGETWVDEPLKFGGGLTKRSYFNKNHVS 439
 411 -----TVMVLKAPPGGFEHAQPDNGTPELFIKGRNFTPDAGVFVYSGDE 454
 440 DDQALMHSSNVNMFETALKLAGDPYTSMSLPNN-----IADAG----- 479
 455 AIMKLNRWYRTRIHSTLTLDNONNVITKARQNKWETGNNDLVLTYNPSYPNLDHORSV 514
 480 --RKLKGLNQVGLKGTIGIDLPNETPGQIEPLTNPNNGYLDLAIQYDVTYTPQLSQVY 537
 515 LFINKKFLV---IDRAIGATGNLGVHWQKESNVPDFKTRVYTYTDRGNLMIQS 571
 538 STIANDGRIOPHIGLSIYESTN-----KDETGPLKRIKIGNVLKNVNSNDEIKEV 589
 572 LNAURTSINEEGKVSVYVYNKELRP-----AFVEE--KPKKNAGTQNFVSIYVP--- 619
 590 QEGFKMAFNEKQG--TGVASFRNTVPSAGKTGTAEVFDQGEPRVNSYTIYGIAPVDDPKLS 648
 620 ----YDQKKAPEISIREKNGNDFEKGKLN 644
 649 PSIVYTNQVPVPPWL---NGDGLGRDVIN 674

RESULT 6
 US-10-369-493-22074
 ; Sequence 22074, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 22074
 ; LENGTH: 532
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-369-493-22074

Query Match 3.5%; Score 123; DB 15; Length 532;
 Best Local Similarity 18.9%; Pred. No. 0.035;
 Matches 109; Conservative 76; Mismatches 179; Indels 214; Gaps 31;
 94 VTREMAKALVHQFQPHKGY-----GYFDYKQINWQW--PVKDNEVRWQLHRVKKWQAM 147
 21 MTNETSDRPLVH--FTPNKGWMDPNGLWYDEKAKWHLFYQYNFNDIVW--GTFPLFWG-- 75

148 ALVYHATGDEKAYREWVYQYSDWARKNPLGLSQNDKFWRPLEVSDRVQSLPTFTSLFV 207
 76 ---HATSD-----LTNW-EDQPIALAPRN----- 97
 208 NSPAPFTAFALMEF-----LNSYHQADYLSHYAEQGNHRLFE 245
 98 DSGAFSGSMVVDYNTSGFNDTIDPRQCVAILTYNTPESEEQVIS--YSLDGGYTTFE 155
 246 AORNLPAGVSPFEK-----SPWRQIGISVLNTEIKQVYAD----- 284
 156 YQNPVLAANSTQPRDPKVPFWYEFESQKIMT--AAKSQDYKIEIYSSDDLKSWKLESAPA 213
 285 ---GMQFEL-----SPIYHVAIDI-----FLKAY 306
 214 NEGLGVQYCEPGLIEVTEQDPSKSYWVMEISINPGAPAGGSNOYFVGSFNGTHFPAP 273
 307 GSAKRVNLEKEFPQSY--VQTVENMIMALISISLPDYNTPMFGSDSWITDKNFRMAQFA-- 362
 274 DNQSRV---VDFGKDYVALQTFEN-----TDPTYGSAL-GIAWAS--NWEYSAFVPT 319
 363 -----SWARVFPANQAIKYFA--TDGKQKAPNLSKALSNAAGFYTFPSGWDKNAIV 412
 320 NPMWSSLSVRKFSLN--TEYQANPETELINLKAEPILN--ISNAG-----PWSRFAIN 369
 413 MVLKASPPGFEHAQPDNGTPELFIKGRNFTPDAGVFVYSGDEAIMK-----LRNRYOTR 467
 370 TLTKANSYNDLSNSTGTLEFEL-----VYAVNTTQTTISKSVFADLSLWFKGLE 419
 468 -----IHSTLTLDNONNVITKARQNKWETG-----NNLDVLTYNPSYPN----- 507
 420 DPEYLRMGFEVSASSPFLDRGNSKVKFVKENPYFTNRMSVNNQPFKSENDLSYKVVYGL 479
 508 LDHORSVLFIN-----KKVELVIDRAIGATGNLGV 538
 480 LDQNLIELYENDGVSTNTYFMTTGNALGSVNNMTGV 517

RESULT 7
 US-10-369-493-23212
 ; Sequence 23212, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 23212
 ; LENGTH: 677
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilis
 US-10-369-493-23212

Query Match 3.4%; Score 120; DB 15; Length 677;
 Best Local Similarity 18.8%; Pred. No. 0.097;
 Matches 123; Conservative 83; Mismatches 216; Indels 232; Gaps 32;
 41 YSGLEKVNKAVAGNYDDAAKALLAYY---REKSKAREPDFSNAEKPADIRQPTDKVTRE 97
 105 FSGSAVVDKNTSGFQTKGKPLVIAIYTDQREGHQVQSIAYSNDK-----GRTWTK 155
 98 MADKALVHQFQPHKGYFDYDGKIDNQWMPVKDNEVRWQLHRVKKWQAMA-----LVYH 152
 156 YAGNPVI-----PNPG-----RDPKVFWEYKEKKWVWVLAAGDRILIY- 197

QY 153 ATGDEKYAREWVYQYSDWARKNPLGLSQNDKFWVRPLEVSDRVQSLPPTSLFVNSPAF 212
Db 198 ---TSKNLKWITY-----ASEFGODQGHGCV-----ECPDLFELPVDGNPN 237
QY 213 TPAFLMEF-----LNSVHQQADLYLTHYAEQG---NHRLF-EAQRNLFAGVSFPPEF 259
Db 238 QKKWQVQVNGVNGVSGGGMQYFVGDGFTGFKNIENPENKVLWTDYGRDFVAAVSWSDI 297
QY 260 --KDSRP-----WROTGISVLNTEIKQVYADGMQFELSPIYHVA 298
Db 298 PSTDSRRLMLGWMNSWQYANDVPTSPWRS--TSIPRELKKAFTEGVRVQVTPVKELET 355
QY 299 I-----DIFLKAYGSAKRVNLE-KEFPQS-----YVQTVENMIM--- 331
Db 356 IRGTSKKKNLITSPASHVNLAQSGDAVEINAEFKVSPGSAABFGFKVKTGENQTKVG 415
QY 332 ----ALISISLPDYNTPMFGDSWIITDKN---FRMAQFASWARVFPANQAIKYFATDGKQ 383
Db 416 YDRNAKLFVDRSESGNDTFNPAFTGKETAPLKPVNGKVKLRI FVDRSSVEVFGNDGKQ 475
QY 384 GKA----PNFLSKAL----SNAG-----FYTPRSGWMDKKNATVMVLK 416
Db 476 VITDIILPDRSGKGLYAAANGVKVKSITIHLKKVWGTTPTFMSNMTCW---TTV----- 528
QY 417 ASPPGEFHAQPNGTPELFIKGRNFTPDAGVFVVS-----GDEA 455
Db 529 -----NGTWADTIEGQKQSGDGSFLLSSASGSDFTYESDIITKONGRGAGA 576
QY 456 IMKLRNWRQTRIHSITLDNQNMVITKARQNKWETGNNLDVLTYPNYPNLDHQRSVL 515
Db 577 LM-----FRSKDKAKNGYLANVDAKHDLVKFKFKFENGASVIAEYKTP-----I 620
QY 516 FINKKYFL-----VIDRAIGEA-----TGNI/GVH-WQKEDSNPVF 550
Db 621 DVNKKYHLKTEABGRFKIYLDRLVIDAHSVFSEGFGLNVW-----DATAVF 670

RESULT 8
US-10-369-493-2315
; Sequence 2315, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2315
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2315

Query Match 3.4%; Score 118.5; DB 15; Length 579;
Best Local Similarity 20.9%; Pred. No. 0.11;
Matches 111; Conservative 75; Mismatches 195; Indels 151; Gaps 31;
QY 112 GYGFDYGDKNQWMPVKD-NEVRWQLHRVQWQAMALVYHATGDEKYAREWVYQYSDW 170
Db 70 GYDVSY-KQIDSRYGTLDELRLMKALHERDMKLVMDLVNITSQ---HEW-FKESRS 124
QY 171 ARKNPLGLSQNDKFWVRPLEVSDRVQSLPPTSLFVNSPAFTPAFLMEFLNSYHQOAY 230
Db 125 SKTNP-----KRDWYFKWPARYNEKGERLPP-----NNWRSYFDT 159

QY 231 LSTHYAEQGNHRLFEAQRNLFAGVSFPPEK-DSPRWRQTGISVLNTEIKQVYADGMQFE 289
Db 160 SAWEDEATQ-----EYLLHWS-VGQPDNLNWTETPKVREAVHDLIRFWLDRGV--DG--PR 210
QY 290 LSPYHVAADIFLKA-----YGSARKRVNLEKEFFQSYQVTVENMIMALI 334
Db 211 LDAINWISKDQRFLOAPITDDRYEYQLAYQYANGPRIH-----EYLANGIGNI----- 258
QY 335 SISLDPYNTMFGD-SWITDKN-----FRMAQFASWARVFPANQAIKYFATD 380
Db 259 ---LTEYDAFSGEMPYVLDITNEILHVVGADRRELTMIPODFVDLDDPNQH-KYIEGS 314
QY 381 GKQKAPNFLSK---ALSNAGF-YTFRSGWMDKKNATV-MVLKASPPGEPHQAQPDNGTTEL 434
Db 315 WELSDKLKSLKQWQALLSGGGWNASFTENHDQTRTVSRYLSDSPKYRAYSKLMALFII 374
QY 435 FTKGRNFTPDAGVFVYSGDEAIMKLRNWRQTRIHSITLDNQNMVITKARQNKWETGNN 494
Db 375 FQSG---TP---FVFGQGE--LALANIPRDPIDEYLDVETQNF-----WK--- 412
QY 495 LDVLYTTPSNYPLNDHQRSVLFIKKYFLVIDRAIGEAATGNLGVHQLKEDSNPVEDTK 554
Db 413 --LFMSGNPSQBEI--EKTMDIVNKR-----ARDNGRTPMEHW---DSSPENGFTK 455
QY 555 -----NRVYTTYRDGNNLMIQSLNADRTSLNEEGKVSY-----VYNKELK 595
Db 456 AGVKPMWRVTNDYKEWN-----AANQVNDPSPYTFWWSKALELRKELK 498

RESULT 9
US-10-369-493-2321
; Sequence 2321, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2321
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2321

Query Match 3.4%; Score 118.5; DB 15; Length 579;
Best Local Similarity 20.9%; Pred. No. 0.11;
Matches 111; Conservative 75; Mismatches 195; Indels 151; Gaps 31;
QY 112 GYGFDYGDKNQWMPVKD-NEVRWQLHRVQWQAMALVYHATGDEKYAREWVYQYSDW 170
Db 70 GYDVSY-KQIDSRYGTLDELRLMKALHERDMKLVMDLVNHTSDQ---HEW-FKESRS 124
QY 171 ARKNPLGLSQNDKFWVRPLEVSDRVQSLPPTSLFVNSPAFTPAFLMEFLNSYHQOAY 230
Db 125 SKTNP-----KRDWYFKWPARYNEKGERLPP-----NNWRSYFDT 159
QY 231 LSTHYAEQGNHRLFEAQRNLFAGVSFPPEK-DSPRWRQTGISVLNTEIKQVYADGMQFE 289
Db 160 SAWEDEATQ-----EYLLHWS-VGQPDNLNWTETPKVREAVHDLIRFWLDRGV--DG--PR 210
QY 290 LSPYHVAADIFLKA-----YGSARKRVNLEKEFFQSYQVTVENMIMALI 334
Db 211 LDAINWISKDQRFLOAPITDDRYEYQLAYQYANGPRIH-----EYLANGIGNI----- 258

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA 034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See file wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 70737

LENGTH: 919

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-10-282-122A-70737

Query Match 3.3%; Score 116; DB 12; Length 919;

Best Local Similarity 18.2%; Pred. No. 0.37;

Matches 124; Conservative 84; Mismatches 216; Indels 258; Gaps 31;

QY 56 YDDAAKALLAYYREKSKAREPDFNSAEKPADIR-----QPIDKVTREMAK 101
 DB 370 YDKANKAVTLLKEDKALLKDLFTHTSYPHDWRTKKVI PRATQWFSINKVRQDILDA 429
 QY 102 ALVHQFQPHKGYGEDYCKDINQMWPVKONEVWQLHRVKKWQAMALVYHA-TGDEKYA 160
 DB 430 IEDTNFK-----VDMGKTRIYN--IRDRG-EWVISRQVWGVPLPVFAENGDIIMT 479
 QY 161 REWYQYQSDWARKNPLGLSQDNDKFWVRPLEVSRVQSLLPTTSLFVNSPAFTAFLEMF 220
 DB 480 KETVNHVAD-----LFEKHGNSINWFEAK---ELLPEGFS-HFGSPN----- 518
 QY 221 LNSYHQADYLSTHYAEQGNHR-LFEAQNLFAVGSFPE---PKDSPRWRQTGISVLNTE 276
 DB 519 -GEFTKETDMDVWFDSGSSHRGVLETRPEL-----SEPADLYFGSQYR-----GWNSS 569
 QY 277 IKGOVADGMQFELSPIYHVAIDIFLKAYGSAKRVNLEKEFPQSYQVTVENMIMALISI 336
 DB 570 ITTAVATRGQ-----APYKFLISHG----- 589
 QY 337 SLPDYNTPMFGDSWITDKNFMQAQFASWARVFPANQAIKYFATIDGKQKQKAPNLSKALSN 396

DB 590 -----FVMDG-EGKK---MSKSLGN 605
 QY 397 AGFYFRSGWCKNATVMVLKASPPGEPHAQPDNGTFELFIKGRNFTPDAGVVFVSGDEAI 456
 DB 606 -----VIV-----PDQVVKQKGAIDARLWVSSTYDLSVRI-----SDEIL 641
 QY 457 MKLRNWRQTR-----IHSTLTLDNQNMWITKARONKWKETGNLNDV 497
 DB 642 KQTSDEVYRKENTLRFMLGNINDFNPDTSIAETNLLVEDRYLLNRLREFTASTINNYEN 701
 QY 498 LTYTN-----PSYPN-----LDQRSVLFINCK-----YFLVIDRAIGBAT 533
 DB 702 FDYLNLYQEVQNFNVELSNFYLDYGDILYIEKKSHKRRSMQTVLYQILVD--MTKLL 759
 QY 534 GNLGVH-----W-----OLKEDSNVFDKTKY-----RVYTYRDGNNLMIQSLN 573
 DB 760 APILVHTAEVWMSHTPHVKEESVHLSDMPKVVDVDEELLEKWKNTFMNLRDDVNRALBOAR 819
 QY 574 ADRTSLNEEGKVSIVYNKELKRPFAVFE-----KPKKNAGTQNFVSIYV 618
 DB 820 NEKVIKSLKAKVIGVSGNESSENTAEFLQOQNDLQQLFIVSQVEYKDKVNDG-----VSY 873
 QY 619 PYDQKAPETISIRENKGNDFEK 640
 DB 874 QYG-----DIHIKHAEGEKCEK 890

RESULT 14
 US-09-802-640-32
 ; Sequence 32, Application US/09802640
 ; Publication No. US20030036057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Braun, Andreas
 ; APPLICANT: Kieyn Patrick
 ; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
 ; FILE REFERENCE: 24736-2048
 ; CURRENT APPLICATION NUMBER: US/09/802,640
 ; CURRENT FILING DATE: 2001-03-09
 ; NUMBER OF SEQ ID NOS: 122
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 32
 ; LENGTH: 4563
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; US-09-802-640-32

Query Match 3.3%; Score 116; DB 10; Length 4563;
 Best Local Similarity 16.3%; Pred. No. 4.4;
 Matches 104; Conservative
 QY 218 MEFLNSY-----HQADY-----LSTHYAEQGNHRLFEA-----QBNL----- 250
 DB 3068 IDFLNNYALFLSPSAQQAQSWQSARFNQYKYNQNFSAQNNENIMEAHVINGEANGLDPLN 3127
 QY 251 -----FAGVSPPEEKDSPRWEQTG----- 269
 DB 3128 IPLTIPMRPLPYITIIITPPLKDFSLWBKTLGKLEFKTTKQSFDSLVSQAQYKKNKHSIT 3187
 QY 270 --ISVLNTEIKQVYADGMQFELSPIYHVAIDIFLKAYGSAK-----RVNLEK---BFPQ 320
 DB 3188 NPLAVLCEFTISQSISFDRHFEKN--RNNALDPVTKSYNETKIKFDKYKAERKSHDELPR 3244
 QY 321 SYQVTVENMIMALISLIPDYNTPMFGDSWITDKNFMQAQF-----ASWAVFPANQAIKY 376
 DB 3245 TF-----QIPGYTPVV-----NVEVSPEFTIEMSAFGVYFPKAVSMPS 3282
 QY 377 FATDGGKQKAPNF-----LSKALSNAGF-YT 401
 DB 3283 FSLIGSDVRVPSVTLLIPSLLEPLVHVPRLKLSLPHKFLCTLSHTIFIPAMGNITVDFS 3342
 QY 402 FRSGWCKNATVMVLKASPPGEPHAQPDNGTFELFIKGRNFTPDAGVVFVSGDEAI----- 456


```

Db 3343 FKSS-----VITLNTNAELFNQSD--IVAHLSSSSSVIDALQYKLEGTTLTRKRG 3392
QY 457 -----MKLNRVQRTRHSTLTLDNQNMVITKARQNKWE-----TGNNLDVL 498
Db 3393 LKATALSLSNKEVEGSHNSTVSLTTKNMEVSVAKTTKABIPILRMNFKQELNGNTKSKP 3452
QY 499 TYTNPSYPNLDHORSVLFINKK-----YFLVIDRAIGEAATGNL----- 536
Db 3453 TVSSMEFFKDFNSSMLYSTAKGAVDHKUSLESLSYFSESSTKGDVKSLSRBSYSGT 3512
QY 537 -----GVH-----WOLKEDGNPVPDKTKNRVITYTYRDG--NNLM 568
Db 3513 IASEANTYLSKSTRSSVKLQGTSKIDDINWLEVKENFAGEATLQRIYLSWEHSTKNHLQ 3572
QY 569 IQSL-----NADRTS---LNEEGKVSYYNKNELKRPAPVFKEP-----KKNACTON----- 612
Db 3573 LEGLFTNGEHTSKATLELSPWMSALVQVHASQPSSEHDFPDLGQBEVALNANTKNQKIR 3632
QY 613 FVSIVVPYDQKAPESIRENKGNDEPKGLNLTTLING 651
Db 3633 WKNEVRHSGFSQSVEL-----SNDQEKALHLDIAGSLLEG 3667

```

RESULT 15

```

US-10-282-122A-48445
; Sequence 48445, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsythe, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48445
; LENGTH: 987
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48445

```

```

Query Match 3.3%; Score 115.5; DB 12; Length 987;
Best Local Similarity 19.1%; Pred. No. 0.46;
Matches 105; Conservative 79; Mismatches 188; Indels 179; Gaps 24;

QY 11 VFAVIALSSGNILAQSSSLTRKDFDHINLEYSGLEKVNKA-----VAAGNYDDAAKALL 64
Db 282 VFSKAAETLGEVTTESDAUTQNALYHLMGLAYLHLAEKNKARMAFEQAASANLKIKEQA 341
QY 65 AY-----YREKSKAREP---DFSNAE-----KPA 85
Db 342 AYNVALCIHETSYSARGESVTFEKLNEFPNSEYAEMVSSYLVEYVMNTRSYEAALKSI 401
QY 86 D-----IROPIDIKVTREMAKALVH-QOPHKGYGYFDYKQKINQWQMPVKDNEVRW 136
Db 402 DRIAHPGKRILEAKQRIPLQGLGTQAFANTQFE--QAIGYFDRSLGLGQYNROTQADALYW 459
QY 137 -----OLHRVKWQWQAMALVYHATGDEKYAREWVYQYSDWARKN 174
Db 460 RGEAYYRLNRMEEAENFTDYLOLTQOTNEMVALAHYINLGYIAFHQKDYTOQNWFRK- 518
QY 175 PLGLSQDNDFVWRPLEVSDRVQSLPPTFS---LFVNSPAFTPAFLMEFLNSYHQADY 230
Db 519 -----YISLEKGENKTALADAYNRIGDCYLDVNFDEA-----KHYYSOAEA 560
QY 231 LSTHYAEQGNHRLFE-----AORNLPAGVSPFPFKDPSRWRQGISVLNT 275
Db 561 MNT---PSGDYSFYQALVSLGQKDYSGKITLINRLAG---KYPASP---YAISSALYE 609
QY 276 EIKKQVYADGMO-----PELSPIYHVAADIEFKAYGSA---KRVNLEKEFPQ 320
Db 610 KGRSYVLMNNOQAIASFKELLAKYEPSPVSKAAAEIGLLYYQNEEDYDOAINAYKOVVQ 669
QY 321 SYVQTVENMIMALISISLDPYNTPMFGDSWITDKNFRMAQFASWARVFPAN-----Q 372
Db 670 KYPGSD-----ARLAMEDLK-----SIYVDMN-RIDEFAALASAMFCNIFPDASEQD 716
QY 373 AIKYFATD-----GKQKAPNFLSKALS-----NAGFYTFRSGWQKNATVMVLKAS 418
Db 717 SLTYMAAEKIYIRGRVEQAKESFGKYLQTFPDGAFGLNAHYLCLIGKEQKNYDMILEHS 776
QY 419 PGCEFHQAPDN 429
Db 777 --GKLLXPDN 785

```

Search completed: March 18, 2004, 04:13:39
Job time : 48 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 20:48:13 ; Search time 29 Seconds
(without alignments)
961.478 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494

Sequence: 1 MTTKFKRIIVFAVIALSSG.....KGLNLTITNGKQLVLVP 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgm2_6/ptodata/1/iaa/5A-COMB.pep:*
2: /cgm2_6/ptodata/1/iaa/5B-COMB.pep:*
3: /cgm2_6/ptodata/1/iaa/6A-COMB.pep:*
4: /cgm2_6/ptodata/1/iaa/6B-COMB.pep:*
5: /cgm2_6/ptodata/1/iaa/PCTUS-COMB.pep:*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3494	100.0	659	1 US-08-258-639A-4	Sequence 4, Appli
2	3494	100.0	659	2 US-08-900-951-4	Sequence 4, Appli
3	3494	100.0	659	5 PCT-US95-07391A-4	Sequence 4, Appli
4	126.5	3.6	698	4 US-09-134-001C-3632	Sequence 3632, Ap
5	124	3.5	23	1 US-08-258-639A-9	Sequence 9, Appli
6	124	3.5	23	2 US-08-900-951-9	Sequence 9, Appli
7	124	3.5	23	5 PCT-US95-07391A-9	Sequence 9, Appli
8	119	3.4	23	1 US-08-258-639A-10	Sequence 10, Appl
9	119	3.4	23	2 US-08-900-951-10	Sequence 10, Appl
10	119	3.4	23	5 PCT-US95-07391A-10	Sequence 10, Appl
11	119	3.4	1138	1 US-07-973-320-2	Sequence 2, Appli
12	119	3.4	1138	1 US-07-973-320-4	Sequence 4, Appli
13	118	3.4	772	1 US-08-258-639A-2	Sequence 2, Appli
14	118	3.4	772	2 US-08-900-951-2	Sequence 2, Appli
15	118	3.4	772	5 PCT-US95-07391A-2	Sequence 2, Appli
16	116	3.3	4536	4 US-09-180-422B-27	Sequence 27, Appl
17	115.5	3.3	1168	1 US-08-620-717A-9	Sequence 9, Appli
18	114.5	3.3	1167	1 US-08-485-568A-6	Sequence 6, Appli
19	114.5	3.3	1167	2 US-08-590-554A-6	Sequence 6, Appli
20	114.5	3.3	1167	2 US-09-184-223-6	Sequence 6, Appli
21	113.5	3.2	1398	1 US-08-750-532-9	Sequence 9, Appli
22	113.5	3.2	1398	3 US-08-894-818B-8	Sequence 8, Appli
23	113.5	3.2	1398	4 US-09-445-472-6	Sequence 6, Appli
24	111	3.2	927	4 US-09-134-001C-4831	Sequence 4831, Ap
25	110	3.1	4563	4 US-09-108-006C-1	Sequence 1, Appli
26	109	3.1	951	3 US-08-816-346-58	Sequence 58, Appl
27	109	3.1	951	3 US-09-335-411-58	Sequence 58, Appl

28 109 3.1 952 2 US-08-788-674-5 Sequence 5, Appli
29 109 3.1 952 3 US-08-816-346-4 Sequence 4, Appli
30 109 3.1 952 3 US-09-335-411-4 Sequence 4, Appli
31 106 3.0 884 4 US-09-328-352-4598 Sequence 4598, Ap
32 106 3.0 945 4 US-09-198-452A-1030 Sequence 1030, Ap
33 106 3.0 1375 3 US-09-210-361-4 Sequence 4, Appli
34 106 3.0 1375 4 US-09-740-274-4 Sequence 4, Appli
35 105 3.0 443 4 US-09-328-352-6322 Sequence 6322, Ap
36 105 3.0 741 4 US-09-252-991A-22440 Sequence 22440, A
37 105 3.0 965 4 US-09-437-277-3 Sequence 3, Appli
38 104 3.0 1849 4 US-08-851-567B-49 Sequence 49, Appl
39 104 3.0 2516 4 US-08-851-567B-47 Sequence 47, Appl
40 103.5 3.0 1183 2 US-08-447-031A-2 Sequence 2, Appli
41 103.5 3.0 1338 2 US-08-728-470-9 Sequence 9, Appli
42 103.5 3.0 1338 3 US-08-719-641-9 Sequence 9, Appli
43 103.5 3.0 1599 2 US-08-617-697-9 Sequence 9, Appli
44 103.5 3.0 10182 4 US-09-134-001C-3159 Sequence 3159, Ap
45 102.5 2.9 618 3 US-09-299-378-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-08-258-639A-4

; Sequence 4, Application US/08258639A

; Patent No. 5681733

; GENERAL INFORMATION:

; APPLICANT: Su, Hongsheng

; APPLICANT: Blain, Francoise

; APPLICANT: Bennett, Clark

; APPLICANT: Gu, Kangfu

; APPLICANT: Zimmermann, Joseph

; APPLICANT: Musil, Roy

; TITLE OF INVENTION: Nucleic Acid Sequences And Expression

; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hale and Dorr

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington, D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/258,639A

; FILING DATE: 10 JUNE 1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Healey, William J.

; REGISTRATION NUMBER: 36,160

; REFERENCE/DOCKET NUMBER: 104385.116

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)942-8400

; TELEFAX: (202)942-8484

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 659 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-258-639A-4

Query Match 100.0%; Score 3494; DB 1; Length 659;

Best Local Similarity 100.0%; Pred. No. 5.8e-308;

Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTKFKRIIVFAVIALSSGNIQAQSSITRKDFDHINLEYSGLEKVNKAAGNYDDAA 60

```
Db 1 MTTKIFKRIIVFAVIALSSGNLAQSSSITRKDFHINLEYSGLEKVNKAAGNYDDAA 60
Qy 61 KALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYK 120
Db 61 KALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYK 120
Qy 121 DINQMWVVKNEVRWQLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGSLQ 180
Db 121 DINQMWVVKNEVRWQLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGSLQ 180
Qy 181 DNDKFVWRPLEVSDRVQSLPTFSLFVNSPAFTPAFLMEFLNSYHQQADYLSTHYAEOGN 240
Db 181 DNDKFVWRPLEVSDRVQSLPTFSLFVNSPAFTPAFLMEFLNSYHQQADYLSTHYAEOGN 240
Qy 241 HRLFEAQNLNLAGVSPFPEKSPRWQRTGISVLNTEIKKQVYADGMQFELSPIYHVAID 300
Db 241 HRLFEAQNLNLAGVSPFPEKSPRWQRTGISVLNTEIKKQVYADGMQFELSPIYHVAID 300
Qy 301 IFLKAYGSAKRVNLEKEFPQSYVQTVENMIMALISISLDPDNTPMFGDSWITDKNFRMAQ 360
Db 301 IFLKAYGSAKRVNLEKEFPQSYVQTVENMIMALISISLDPDNTPMFGDSWITDKNFRMAQ 360
Qy 361 FASWARVPPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSGWKNATVMVLKASPP 420
Db 361 FASWARVPPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSGWKNATVMVLKASPP 420
Qy 421 GEHQAQPDNGTFFELFIKGRNFTPDAGVYVSGDEAIKMLRNWYQRTIHSHTLTDNQNV 480
Db 421 GEHQAQPDNGTFFELFIKGRNFTPDAGVYVSGDEAIKMLRNWYQRTIHSHTLTDNQNV 480
Qy 481 ITKARQNKWETGNLNLVLTYPNPSYVNLNDRHORSVLFINKYFLVIDRAIGATGNLGVHM 540
Db 481 ITKARQNKWETGNLNLVLTYPNPSYVNLNDRHORSVLFINKYFLVIDRAIGATGNLGVHM 540
Qy 541 QLKEDSNPVDKTKNRVYTTVYRDGNLMIQSLNADRTSLNEEGKVSYYNKKELKRPV 600
Db 541 QLKEDSNPVDKTKNRVYTTVYRDGNLMIQSLNADRTSLNEEGKVSYYNKKELKRPV 600
Qy 601 FEKPKNAGTQNFVSIYVPYDGQKAPESITRENKNDPEKGNLTITINGKQQLVLP 659
Db 601 FEKPKNAGTQNFVSIYVPYDGQKAPESITRENKNDPEKGNLTITINGKQQLVLP 659

RESULT 2
US-08-900-951-4
; Sequence 4, Application US/08900951
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,951
; FILING DATE:
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,639
; FILING DATE: 10 JUNE 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942 8400
; TELEFAX: (202)942 8484
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-900-951-4

Query Match 100.0%; Score 3494; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 5.8e-308;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTTKIFKRIIVFAVIALSSGNLAQSSSITRKDFHINLEYSGLEKVNKAAGNYDDAA 60
Db 1 MTTKIFKRIIVFAVIALSSGNLAQSSSITRKDFHINLEYSGLEKVNKAAGNYDDAA 60
Qy 61 KALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYK 120
Db 61 KALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYK 120
Qy 121 DINQMWVVKNEVRWQLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGSLQ 180
Db 121 DINQMWVVKNEVRWQLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGSLQ 180
Qy 181 DNDKFVWRPLEVSDRVQSLPTFSLFVNSPAFTPAFLMEFLNSYHQQADYLSTHYAEOGN 240
Db 181 DNDKFVWRPLEVSDRVQSLPTFSLFVNSPAFTPAFLMEFLNSYHQQADYLSTHYAEOGN 240
Qy 241 HRLFEAQNLNLAGVSPFPEKSPRWQRTGISVLNTEIKKQVYADGMQFELSPIYHVAID 300
Db 241 HRLFEAQNLNLAGVSPFPEKSPRWQRTGISVLNTEIKKQVYADGMQFELSPIYHVAID 300
Qy 301 IFLKAYGSAKRVNLEKEFPQSYVQTVENMIMALISISLDPDNTPMFGDSWITDKNFRMAQ 360
Db 301 IFLKAYGSAKRVNLEKEFPQSYVQTVENMIMALISISLDPDNTPMFGDSWITDKNFRMAQ 360
Qy 361 FASWARVPPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSGWKNATVMVLKASPP 420
Db 361 FASWARVPPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSGWKNATVMVLKASPP 420
Qy 421 GEHQAQPDNGTFFELFIKGRNFTPDAGVYVSGDEAIKMLRNWYQRTIHSHTLTDNQNV 480
Db 421 GEHQAQPDNGTFFELFIKGRNFTPDAGVYVSGDEAIKMLRNWYQRTIHSHTLTDNQNV 480
Qy 481 ITKARQNKWETGNLNLVLTYPNPSYVNLNDRHORSVLFINKYFLVIDRAIGATGNLGVHM 540
Db 481 ITKARQNKWETGNLNLVLTYPNPSYVNLNDRHORSVLFINKYFLVIDRAIGATGNLGVHM 540
Qy 541 QLKEDSNPVDKTKNRVYTTVYRDGNLMIQSLNADRTSLNEEGKVSYYNKKELKRPV 600
Db 541 QLKEDSNPVDKTKNRVYTTVYRDGNLMIQSLNADRTSLNEEGKVSYYNKKELKRPV 600
Qy 601 FEKPKNAGTQNFVSIYVPYDGQKAPESITRENKNDPEKGNLTITINGKQQLVLP 659
Db 601 FEKPKNAGTQNFVSIYVPYDGQKAPESITRENKNDPEKGNLTITINGKQQLVLP 659

RESULT 3
PCT-US95-07391A-4
; Sequence 4, Application PC/TUS9507391A
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
```

APPLICANT: ZIMMERMANN, Joseph
TITLE OF INVENTION: Nucleic Acid Sequences And Expression
TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
TITLE OF INVENTION: Flavobacterium heparinum
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07391A
FILING DATE: 09-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,639
FILING DATE: 10 JUNE 1994
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 104385.116PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-07391A-4

Query Match 100.0%; Score 3494; DB 5; Length 659;
Best Local Similarity 100.0%; Pred. No. 5.8e-308;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTKIFKRIIVFAVIALSSGNILAAQSSITRKDFHINLEYSGLEKVNKAVAGNVDAA 60
DB 1 MTTKIFKRIIVFAVIALSSGNILAAQSSITRKDFHINLEYSGLEKVNKAVAGNVDAA 60

QY 61 KALLAYREKSKAREPDFSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYPDYGK 120
DB 61 KALLAYREKSKAREPDFSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYPDYGK 120

QY 121 DINWQMPVKDNEVRWOLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGLSQ 180
DB 121 DINWQMPVKDNEVRWOLHRVKKWQAMALVYHATGDEKYAREWVYQYSDWARKNPLGLSQ 180

QY 181 DNDKFWREPLEVSDRVQSLPPTFSLFVNSPAFTAPFIMEFLNSYHQADYLSTHYAEQGN 240
DB 181 DNDKFWREPLEVSDRVQSLPPTFSLFVNSPAFTAPFIMEFLNSYHQADYLSTHYAEQGN 240

QY 241 HRLFEAQRNLFAGVSPPEPKDSPRWQTGIVLNTGKQVADGCMQFELSPIYHVAID 300
DB 241 HRLFEAQRNLFAGVSPPEPKDSPRWQTGIVLNTGKQVADGCMQFELSPIYHVAID 300

QY 301 IFLKAYGSAKRVNLEKEFFQSVQVTVENMIMALISLSLPDYNTPMGDSWITDKNPRMAQ 360
DB 301 IFLKAYGSAKRVNLEKEFFQSVQVTVENMIMALISLSLPDYNTPMGDSWITDKNPRMAQ 360

QY 361 FASWARVFPANQAIKYFATDGGKQKAPNLSKALSNAAGYTFRSGWDKNATVMVLKASPP 420
DB 361 FASWARVFPANQAIKYFATDGGKQKAPNLSKALSNAAGYTFRSGWDKNATVMVLKASPP 420

QY 421 GEFHAQPDNGTTELFKIGNRFTPDAGVFVYSGDEAIMKLRNRYQTRIHSTILTLDNQNV 480
DB 421 GEFHAQPDNGTTELFKIGNRFTPDAGVFVYSGDEAIMKLRNRYQTRIHSTILTLDNQNV 480

QY 481 ITKARQNKWETGNLNDVLTYTNPSPNLDHORSVLINKKYFLVIDRAIGEATGNLGVHW 540
DB 481 ITKARQNKWETGNLNDVLTYTNPSPNLDHORSVLINKKYFLVIDRAIGEATGNLGVHW 540

QY 541 QLKEDSNPVFDKTKNRVYTYRDGNLMIQSLNADRTSLNEBEGKVSYYVYNKELKRPVF 600
DB 541 QLKEDSNPVFDKTKNRVYTYRDGNLMIQSLNADRTSLNEBEGKVSYYVYNKELKRPVF 600

QY 601 FEKPKNAGTQNFVSIYVYDQKAPETISIRENKGNDFEKGKLNLTTLTINGKQOLVLP 659
DB 601 FEKPKNAGTQNFVSIYVYDQKAPETISIRENKGNDFEKGKLNLTTLTINGKQOLVLP 659

RESULT 4
US-09-134-001C-3632
; Sequence 3632, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3632
; LENGTH: 698
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3632

Query Match 3.6%; Score 126.5; DB 4; Length 698;
Best Local Similarity 18.6%; Pred. No. 0.013;
Matches 139; Conservative 109; Mismatches 284; Indels 217; Gaps 31;

QY 9 ITVFAVIALSSGNI-LAQSSITR--KDFDHINLEYSGLEKVNKAVAGN-----Y 56
DB 32 VFIFAIVLVRLGYLQAQSHYKQLIKNDENIT-----VNESVPRGILDRNGKVLV 83

QY 57 DDAAKALLAYREKSKAREPDFSNAEKPAD-IRQPIDKVTREMAKALVHQFQPHKGYP 115
DB 84 DNASKMSITYTRNRKTSQKEMLTAKLTLTKMDTKITER--DK-----127

QY 116 FDYKGDINWQMP-----VKDNEVRWOLHRV--KWWQA 146
DB 128 ----KDFWLTQMPSPAKLMRKEQLMLEDGSISQDFDTQLDRDKTKGKQLTKDLOV 183

QY 147 MALV-----YHATGDEKYAREWVYQYSDWARKNPLGLSQDND 183
DB 184 LAIYREMNAGSTLDPOTIKNEOVSEKEYAAVSQQLSKLPGVNTTMDWDRKYPVG--DTL 240

QY 184 KFWWREPLEVSDRVQSLPPTFSLFVNSPAFTAPFIMEFLNSYHQADYLSTHYAEQGNHRL 243
DB 241 RGIFG--DVYSTEGIPKELT-----EQYLSKGYSRNDRVGSYLEYQYEDV 285

QY 244 FEAQRNLFAGVSPPEPKDSPRWQTGIVLNTGKQVADGCMQFELSPIYHVAID 303
DB 286 LKGTQKQ-----KYTTDKSGRVISSEVLNPGSR-----GHDQLT-----IDIDL 326

QY 304 KAYGSAKRVN--LEKEFFQSVQVTVENMIMALISLSLPDYNTPMGDSWITDKNPRMAQ 361
DB 327 Q-----KKVESLEKEQISKLSQGAQKMDNALVMVQNPKNGDILAIAGKQIDKQGLKDY 381

QY 362 --ASWARVFPANQAIK-YFATDGGKQKAPN-----FLSKALSNAAGYTFRSGWDKNA 410
DB 382 DIGNFTAQTYTVSSVGGTLLAGYQNKAINVGETVMVDEPLKFPQGGLTFRSYFNKGNHVS 441

QY 411 -----TVMVLKASPPGEFHAQPDNGTTELFKIGNRFTPDAGVFVYSGDE 454

Db 442 DDQALMHSSNVYMEFKTAGDPTSCMSLPNN-----IADAG----- 481
QY 455 AIMKLRNRYRQTHSTTLDNONVITKARONKWETGNNLDVLTYPSPYNLHDQRSV 514
Db 482 --RKLKGLNQGVLGKTKGIDLPNETPGQIEPLTNNGNYLDAIGQDVTYPLQLSQVV 539
QY 515 LFINKKYFLV---IDRAEGEATGNLGVHWQKEDSNPVFDKTKNRVYTYRDNLMIQS 571
Db 540 STIANDGTRIOPHIGLSIYESTN-----KDETPLRKRTKGNVLKNNVNSNDIEKEV 591
QY 572 LMADETSNEEBGKVSYYNKLKRP-----AFVFE--KPKKNAGTQNFVSIYVP--- 619
Db 592 QEGPKWAFNEKOG-TGYASFRNTVVPVSAGKTAEVFDQGEPRVNSTVIGYAPVDDPKLS 650
QY 620 ----YDGKAPBEISIRENKNDPEKGLN 644
Db 651 FSIYVTNQVPFPPWL---NGGDLGRDVIN 676

RESULT 5
US-08-258-639A-9
; Sequence 9, Application US/08258639A
; Patent No. 5681733
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,639A
; FILING DATE: 10 JUNE 1994

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELEPHONE: (202)942-8400
; TELEFAX: (202)942-8494
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-258-639A-9

Query Match 3.5%; Score 124; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 VLKASPPGGEFFHAQPDNGTFELFI 436
Db 1 VLKASPPGGEFFHAQPDNGTFELFI 23

RESULT 6
US-08-900-951-9
; Sequence 9, Application US/08900951
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,951
; FILING DATE:

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,639
; FILING DATE: 10 JUNE 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942,8400
; TELEFAX: (202)942,8484

INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-900-951-9

Query Match 3.5%; Score 124; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 VLKASPPGGEFFHAQPDNGTFELFI 436
Db 1 VLKASPPGGEFFHAQPDNGTFELFI 23

RESULT 7
PCT-US95-07391A-9
; Sequence 9, Application PC/TUS9507391A
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
; APPLICANT: ZIMMERMANN, Joseph
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.

ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07391A
FILING DATE: 09-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,639
FILING DATE: 10 JUNE 1994
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 104385.116PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-07391A-9

Query Match 3.5%; Score 124; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 VLKASPPGCFEHAQPDNGTFELPI 436
DB 1 VLKASPPGCFEHAQPDNGTFELPI 23

RESULT 8
US-08-258-639A-10
; Sequence 10, Application US/08258639A
; Patent No. 5681733
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,639A
FILING DATE: 10 JUNE 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Healey, William J.
REGISTRATION NUMBER: 36,160
REFERENCE/DOCKET NUMBER: 104385.116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400

TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-258-639A-10

Query Match 3.4%; Score 119; DB 1; Length 23;
Best Local Similarity 91.3%; Pred. No. 0.00022;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 KALVHQFQPHKGYGYFDYDKDIN 123
DB 1 KALVHFWFPHKGYGYFDYDKDIN 23

RESULT 9
US-08-900-951-10
; Sequence 10, Application US/089000951
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,951
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,639
FILING DATE: 10 JUNE 1994
ATTORNEY/AGENT INFORMATION:
NAME: Healey, William J.
REGISTRATION NUMBER: 36,160
REFERENCE/DOCKET NUMBER: 104385.116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-900-951-10

Query Match 3.4%; Score 119; DB 2; Length 23;
Best Local Similarity 91.3%; Pred. No. 0.00022;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 KALVHQFQPHKGYGYFDYDKDIN 123
DB 1 KALVHFWFPHKGYGYFDYDKDIN 23

RESULT 10

PCT-US95-07391A-10
; Sequence 10, Application PC/TUS9507391A
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
; APPLICANT: ZIMMERMANN, Joseph
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07391A
FILING DATE: 09-JUNE-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/258,639
FILING DATE: 10 JUNE 1994
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 104385.116PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484

INFORMATION FOR SEQ ID NO:

10:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US95-07391A-10

Query Match 3.4%; Score 119; DB 5; Length 23;
Best Local Similarity 91.3%; Pred. No. 0.00022;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 101 KALVHQFQPHKGYGYFDYGDKN 123
Db 1 KALVHWFPHKGYGYFDYGDKN 23

RESULT 11

US-07-973-320-2
; Sequence 2, Application US/07973320
; Patent No. 5286486

GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Fu, Jenny M.
TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene
TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/973,320
FILING DATE: 19921106
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/788,638
FILING DATE: 6-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA68.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1138 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: dakota
INDIVIDUAL ISOLATE: HDS11
IMMEDIATE SOURCE:
LIBRARY: Lamdagem (TM)-11 library of J.M. Fu
CLONE: 511
US-07-973-320-2

Query Match 3.4%; Score 119; DB 1; Length 1138;

Best Local Similarity 19.8%; Pred. No. 0.14;

Matches 132; Conservative 85; Mismatches 206; Indels 244; Gaps 39;

Qy 48 NKAVA-----AGNYDDAKALLAYREKSKAREPDSNAEKPADINQPIDKVTREMADKAL 103
Db 121 NKALAELEGLNN-----LTIYQQ-----ALEDLNPNDDPATITRVIDRF--RILD-AL 167

Qy 104 VHQFQPHKGYGYFD-----YGDKNMQM-----WPVKDNEV----- 134

Db 168 PESYMPSPFRVAGYEIPLLTVYQAANLHALLRSTLYGDKWGTQNNIENYRQKXHI 227

Qy 135 -RWOLHRVKKWQA-MALVYHATGDE-----KYAEWVYQYSDWAR-----KNPLGLSQDND 183

Db 228 SEYSNHCWKYNSGLSRLNGSTYEQWYNRFRREMILMVLIDIAAVPIYDPRMYSMETS 287

Qy 184 KFWWR-----PLEVSDRVQSLPPTFSLFVNSPATPAFLMEFLNSYHQOADYLSTHYAEQ 238

Db 288 TQLTREYVTDPISLISNPDPDGPSQMENTAFRTPHLV-----DYLDSELYTYT 336

Qy 239 GNHELF-EAQRNLF-----AGVSPFEKDSRWTGTISVLNTEIKQVYADGMQFELSP 292

Db 337 SKYKAFSHEIQPDLFYWCVHKVSPKSEQSNLY-TTGT-----YKTSGYISSG 384

Qy 293 IYHVAADIF-----LKAYSKAKRVNLEKEFPQSYQVQTVENMIM 331

Db 385 AYSFRGNDIVRTLAAPSVWVYPTQNYGVGEQVEFYGVKGVHVRGD--NKYDLTYDSIDQ 442

Qy 332 -----ALISLSLPDNYT---PMIGDSWITDKNFRMAQFASWAVFP 369

Db 443 LPDGEPIHEKYHRLCHATAISKSTPDYDNATIPF--SW-----THRSABY--YNRILYP 494

Qy 370 AN-----QAIKYFATDGKQG--KAPNFLSKALSNAAGFYFRSGW--DKNATVWLKASPPG 421

Db 495 NKIKKIPAVKMYKLDLSTVVKGPGFTGGDLVKRG-----SNGYIGDIKATV-----NSPLS 546

Qy 422 E-----FHAQPDNGTGFELFIKGRNFTPDAGVFVYSGDEAIMKLNRNMYRQTRIHSHTLTD 475

Db 547 QKRVVRVRYATSVSGLNFVFI----- 567
QY 476 NQNMVITKARONKWT--GNLDDVLT--TNSYPN-----LDH--QRSVL 515
Db 568 NDEIALQKNFQSTVETIGEGKD-LTYGSGYIEYTTTQFPNEHPKITLHLNHLNNSPF 626
QY 516 FINKKYFLVIDRAIGEATGNLGVHOLKEDSNPVDKTKNRVYTTVRDGNLNL--IQSL 572
Db 627 YVDSIEFIPVD-----VNYDEK-----LEKAKAVNTLFTEGRNALQKVVTDY 671
QY 573 NADRTSL 579
Db 672 KVDQVSI 678

RESULT 12

US-07-973-320-4
; Sequence 4, Application US/07973320
; Patent No. 5286486
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Fu, Jenny M.
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/973,320
; FILING DATE: 19921106
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/788,638
; FILING DATE: 6-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA68.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1138 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: kumamotoensis
; INDIVIDUAL ISOLATE: HD867
; IMMEDIATE SOURCE:
; LIBRARY: Lamdagem (TM)-11 library of J.M. Fu
; CLONE: 867
US-07-973-320-4

Query Match 3.4%; Score 119; DB 1; Length 1138;
Best Local Similarity 19.8%; Pred. No. 0.14;
Matches 132; Conservative 85; Mismatches 206; Indels 244; Gaps 39;

QY 48 NKAVA-----AGNYDDAKALLAYREKSKAREPDSFNAEKPADIRQIPDKVTREMAKAL 103
Db 121 NKALAELEGLNN-----LTIYQQ-----ALEDWLNNEPDDPATITRVIDRF--RIID-AL 167
QY 104 VHQFQPHKGYGYFD-----YCKDINWQM-----WPKVDNEV----- 134
Db 168 PESYMFSPVAGYEIPLLTVAQAANLHALLROSTLYGDKWGFTQNNIENYNEQKKHI 227
QY 135 -RWQLHRVKKWQA-MALVYHATGE-----KYAREWVYQYSDWAR-----KNPLGLSQDND 183
Db 228 SEYSNHCWYNSGLSRLNGSTYEQWINVYPREREMILMVLIDIAAVFPIYDPRMYSMETS 287
QY 184 KEVWR-----PLEVSDRVQSLPPTESLFVNSPAFTAFMEFLNSVHOQADVLSTHYAQ 238
Db 288 TOLTREVYTDPLSLISNPDIGPFSQMENTAFRPHLV-----DYLDLYIYT 336
QY 239 GNHRLF--EAORNLF---AGVSFPEFKDSPRWOTGISVLNTEIKKQVYADGMOFELSP 292
Db 337 SKYKAFSHEIQPDLFWCVHKVSPFKKSQSNLY-TTGI-----YGKTSYVSSG 384
QY 293 IYHVAADIF-----LKYGSAKRVNLEKEPPOSYVQTVENMIM 331
Db 385 AYSFEGNDIYRTLAAPSVVVYPTQNYGVEQVEFYGVKGVHYRGD--NKYDLTYDSIDQ 442
QY 332 -----ALISISLPDYN-----PMFGDSWITDKNFMQAQFASWARVFP 369
Db 443 LPPDGEPIHEKYTHRLCHATAISKSTPDYDNATIF--SW-----THRSAEY--YNRIVP 494
QY 370 AN-----QAIKYPATDGKQ--KAPNFLSKALSNAGFYTFRSGW--DKNATVMVLKASPPG 421
Db 495 NKIKKIPAKVMYKLDLSTVVKGPGFTGDLVKRG---SNGYIGDIKATV---NSPLS 546
QY 422 E-----FHAQPDNGTPELFKGRNFTPDAGVVFVSGDEAIMKLRNWRQTRIHTLTLD 475
Db 547 QKRVVRVRYATSVSGLNFVFI----- 567
QY 476 NQNMVITKARONKWT--GNLDDVLT--TNSYPN-----LDH--QRSVL 515
Db 568 NDEIALQKNFQSTVETIGEGKD-LTYGSGYIEYTTTQFPNEHPKITLHLNHLNNSPF 626
QY 516 FINKKYFLVIDRAIGEATGNLGVHOLKEDSNPVDKTKNRVYTTVRDGNLNL--IQSL 572
Db 627 YVDSIEFIPVD-----VNYDEK-----LEKAKAVNTLFTEGRNALQKVVTDY 671
QY 573 NADRTSL 579
Db 672 KVDQVSI 678

RESULT 13

US-08-258-639A-2
; Sequence 2, Application US/08258639A
; Patent No. 5681733
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/258,639A
APPLICATION NUMBER: US/08/258,639A
FILING DATE: 10 JUNE 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Healey, William J.
REGISTRATION NUMBER: 36,160
REFERENCE/DOCKET NUMBER: 104385.116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942-8400
TELEFAX: (202)942-8484
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 772 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-258-639A-2

Query Match 3.4%; Score 118; DB 1; Length 772;
Best Local Similarity 20.5%; Pred. No. 0.09;
Matches 128; Conservative 79; Mismatches 229; Indels 188; Gaps 30;
QY 145 QAMALVYHATGDEKYAREWVYQYSD---WARKNPLG-LSQNDKFWVRPLEVSDRV--QS 198
DB 109 ELMALNYLMTYDKPKVGREAITSIIDTETATFKPAGDISRGIGLFWMTGAIVYDWCYDQL 168
QY 199 LPPTFSLFVNS-----PAFTPAFLMEFLNSYHQOQADYLSHYAEQGNHRLFEA 246
DB 169 KPEKTRFVKAFVRLAKMLEGYPVKDKSVGHASEMMIMRDLSSVGLAIYDE---PPE 225
QY 247 QRNLFAGVSPPEFKDSPRW-----RQTGISVLNTEIKKQVYA-----DGM 286
DB 226 MYNLAAGRFKHEHLVARNWFYPSSHNYHQGMSYLVNFTNDLFWLDRMGAGNVFNPQG 285
QY 287 QFELSPYHVAADIFLKAVGSAKRVNLEKEFPQSYQTVENMIMALLISISLPDYNTPMF 346
DB 286 QFILDYAIKRRPDQIILAGD---VDYSRKPKYYT-----MPALLAGSY--YKDEYL 334
QY 347 GDSWITDKNFRMAOFASWARVFPANQAIKYFATPKQG-KAPNFLSKAL---SNAGFYTF 402
DB 335 NYEFLKDPN-----VEPHCKLFEFLWRDTQLGSRKRPDDLPLSRYSGPSFGWMIA 383
QY 403 RSGWDKNATVMVLKASPPGEF---HAQPDNGTFELFKGRNFTPDAGVFV-YSGDEAIM 457
DB 384 RTGWPESVIAEMKVN---EYSFLNHQHDAGAFQIYYKG-PLAIDAGSYTGSSGGYNP 439
QY 458 KLRNRYQTRIHSTLTL-----DNQNMVITKARQ-----486
DB 440 HNKYFFKRTIAHNSILLIYDPKETSSGSGSDHTDFAANDGGQELPKGKIAPRDLKEM 499
QY 487 --NKWETGNL-----DVLTYTPNSYPL-----DHORSVLFINXK-----520
DB 500 LAGDFRTGKILAQGFDPNQ---PDYTYLKGDITAAVSAKVKEKRSFLEFLNLKDAKVP 556
QY 521 -YELVIDRAIGEATGNLGVHQLKEDSNPVFDK-----TKNRVITYYTDGNN--566
DB 557 AAMIVFDKVA-----SNPDPFKFWLHLSIEQPEIKGNQIITIKRTKNGDS 601
QY 567 -----LMIQSLNADRTSLNEEBGKVSYYV-----NKLKRPAPVFE-KPKK 606
DB 602 GMLVNTALLPDAANSNITSIG-GRGKDFWVGTVNTNDPKFGTDEALERGERWVEITPKK 660
QY 607 NAGTQNFVSIVPYDG--QKAPBI 628
DB 661 AAAEDYILNVIQIADNTQOKLHEV 684

RESULT 14

US-08-900-951-2

; Sequence 2, Application US/08900951

Patent No. 5919693
GENERAL INFORMATION:
APPLICANT: Su, Hongsheng
APPLICANT: Blain, Francoise
APPLICANT: Bennett, Clark
APPLICANT: Gu, Kangfu
APPLICANT: Zimmermann, Joseph
APPLICANT: Musil, Roy
TITLE OF INVENTION: Nucleic Acid Sequences And Expression
TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
TITLE OF INVENTION: Flavobacterium heparinum
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hale and Dorr
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington, D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,951
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/258,639
FILING DATE: 10 JUNE 1994
ATTORNEY/AGENT INFORMATION:
NAME: Healey, William J.
REGISTRATION NUMBER: 36,160
REFERENCE/DOCKET NUMBER: 104385.116
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)942 8400
TELEFAX: (202)942 8484
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 772 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-900-951-2

Query Match 3.4%; Score 118; DB 2; Length 772;
Best Local Similarity 20.5%; Pred. No. 0.09;
Matches 128; Conservative 79; Mismatches 229; Indels 188; Gaps 30;
QY 145 QAMALVYHATGDEKYAREWVYQYSD---WARKNPLG-LSQNDKFWVRPLEVSDRV--QS 198
DB 109 ELMALNYLMTYDKPKVGREAITSIIDTETATFKPAGDISRGIGLFWMTGAIVYDWCYDQL 168
QY 199 LPPTFSLFVNS-----PAFTPAFLMEFLNSYHQOQADYLSHYAEQGNHRLFEA 246
DB 169 KPEKTRFVKAFVRLAKMLEGYPVKDKSVGHASEMMIMRDLSSVGLAIYDE---PPE 225
QY 247 QRNLFAGVSPPEFKDSPRW-----RQTGISVLNTEIKKQVYA-----DGM 286
DB 226 MYNLAAGRFKHEHLVARNWFYPSSHNYHQGMSYLVNFTNDLFWLDRMGAGNVFNPQG 285
QY 287 QFELSPYHVAADIFLKAVGSAKRVNLEKEFPQSYQTVENMIMALLISISLPDYNTPMF 346
DB 286 QFILDYAIKRRPDQIILAGD---VDYSRKPKYYT-----MPALLAGSY--YKDEYL 334
QY 347 GDSWITDKNFRMAOFASWARVFPANQAIKYFATPKQG-KAPNFLSKAL---SNAGFYTF 402
DB 335 NYEFLKDPN-----VEPHCKLFEFLWRDTQLGSRKRPDDLPLSRYSGPSFGWMIA 383
QY 403 RSGWDKNATVMVLKASPPGEF---HAQPDNGTFELFKGRNFTPDAGVFV-YSGDEAIM 457
DB 384 RTGWPESVIAEMKVN---EYSFLNHQHDAGAFQIYYKG-PLAIDAGSYTGSSGGYNP 439

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 28, 2003, 20:30:33 ; Search time 85 Seconds
(without alignments)
1230.598 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494

Sequence: 1 MTKIKFKRIIVFAVIALSSG.....KGKLNLTITNGKQQLVLP 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3494	100.0	659	17	AAR89265
2	250.5	7.2	634	23	ABP26592
3	231	6.6	641	23	ABP26593
4	176	5.0	546	24	ABU00675
5	137.5	3.9	1165	23	ABP29975
6	137.5	3.9	1174	23	ABP26712
7	128	3.7	25	22	ABB06932
8	126.5	3.6	698	23	ABP38787
9	124	3.5	23	17	AAR89270
					Heparinase-III-der
					Heparinase-III. F
					Streptococcus poly
					Streptococcus poly
					S. pneumoniae type
					Streptococcus poly
					Streptococcus poly
					Flavobacterium hep
					Staphylococcus epi
					Heparinase-III-der

10	124	3.5	772	21	AAV70201	Mutant F. heparinu
11	123	3.5	532	5	AP40020	Saccharomyces cere
12	123	3.5	532	12	AAV70145	Preinvertease, Sac
13	120	3.4	772	21	AAV70170	Modified F. hepari
14	119	3.4	23	17	AAR89271	Heparinase-III-der
15	119	3.4	772	21	AAV70171	Modified F. hepari
16	119	3.4	772	21	AAV70172	Modified F. hepari
17	119	3.4	772	21	AAV70173	Modified F. hepari
18	119	3.4	772	21	AAV70199	Mutant F. heparinu
19	119	3.4	772	21	AAV70202	Mutant F. heparinu
20	119	3.4	772	21	AAV70208	Mutant F. heparinu
21	118.5	3.4	1138	14	AAR37213	B. t. toxin HD511.
22	118.5	3.4	772	21	AAV70196	Mutant F. heparinu
23	118	3.4	772	17	AAR89264	Heparinase-II. Fl
24	118	3.4	772	21	AAV70158	F. heparinum hepar
25	118	3.4	772	21	AAV70198	Mutant F. heparinu
26	118	3.4	772	21	AAV70204	Mutant F. heparinu
27	118	3.4	772	21	AAV70207	Mutant F. heparinu
28	117.5	3.4	1484	18	AAW55686	H. pylori ORF 07ee
29	117.5	3.4	2902	22	AB46351	H. pylori HPN165 p
30	117	3.3	772	21	AAV70197	Mutant F. heparinu
31	117	3.3	772	21	AAV70200	Mutant F. heparinu
32	117	3.3	772	21	AAV70203	Mutant F. heparinu
33	117	3.3	772	21	AAV70206	Mutant F. heparinu
34	117	3.3	1250	21	AAV91279	Group B Streptococ
35	116.5	3.3	1138	15	AAR46225	Bacillus thuringie
36	116	3.3	4536	19	AAW41262	Apolipoprotein B-1
37	116	3.3	4536	20	AAW96826	Amino acid sequenc
38	116	3.3	4560	23	AAU98981	Human apolipoprote
39	116	3.3	4563	23	AAO15893	Human apolipoprote
40	115.5	3.3	766	20	AAV06426	Avian topoisomeras
41	115.5	3.3	1168	18	AAW16326	Nematocidal toxin
42	115	3.3	1138	15	AAR46226	Bacillus thuringie
43	115	3.3	1252	23	ABP26711	Streptococcus poly
44	115	3.3	1252	23	ABP29749	Streptococcus poly
45	115	3.3	1280	24	ABU00615	S. pneumoniae type

ALIGNMENTS

RESULT 1
AAR89265
ID AAR89265 standard; Protein; 659 AA.
XX AAR89265;
AC AAR89265;
XX
DT 07-APR-1996 (first entry)
XX
DE Heparinase-III.
XX
KW Heparinase-III; heparan sulphate degradation; PCR; primer; cloning;
KW Escherichia coli; polymerase chain reaction; ribosome binding site;
KW toxicity; vector; haemostatic; blood-clotting; antibody;
KW affinity chromatography.
XX
OS Flavobacterium heparinum.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT Peptide /note= "Signal peptide"
FT Peptide 101..123
FT Peptide /note= "Peptide 3C (AAR89271)"
FT Peptide 128..142
FT Peptide /note= "Peptide 3A (AAR89269)"
FT Peptide 414..436
FT Peptide /note= "Peptide 3B (AAR89270)"
XX
PN WO9534635-A1.
XX
PD 21-DEC-1995.
XX
PF 09-JUN-1995; 95WO-US07391.

```
XX PR 10-JUN-1994; 94US-0258639.
XX PA (IBEX-) IBEX TECHNOLOGIES.
XX PA (ZIMM/) ZIMMERMANN J.
XX FI Bennett C, Blain F, Gu K, Musil R, Su H, Zimmermann J;
XX DR WPI; 1996-097381/10.
XX DR N-PSDB; AAQ99227.
XX PS
XX Nucleic acids encoding Flavobacterium heparinum heparinase II and
PT III - for degrading heparin and heparan sulphate, also related host
PT cells, proteins and antibodies, useful in heparinase purificn.
XX PS
XX Claim 16; Fig 9; 75pp; English.
XX
XX The sequence corresponds to Flavobacterium heparinum heparinase-III
CC (EC-4.2.2.8), which degrades heparan sulphate. The sequence is
CC isolated by polymerase chain reaction amplification using e.g.
CC primers 3-1 to 3-4 (AAQ99234-099237) or degenerate primers 3-5 to 3-8
CC (AAQ99238-099241), which are designed based on the sequences of
CC peptides 3A-3C. Toxicity of the genes and natural selection of
CC the host against clones with the entire sequence has been
CC circumvented by cloning sections of the gene separately in
CC Escherichia coli using a vector with a modified ribosome binding
CC site, which increases expression levels. The heparinase-III may be
CC used to neutralise anticoagulant activity. Antibodies against the
CC protein may be used to differentiate between native and recombinant
CC enzymes, and when immobilised they may be used for heparinase
CC purification by affinity chromatography.
XX PS
XX Sequence 659 AA;
XX
XX Query Match 100.0%; Score 3494; DB 17; Length 659;
XX Best Local Similarity 100.0%; Pred. No. 4.8e-279;
XX Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MTTKIFKRIIVFAVIALSSGNILAQSSITRKDPFHINLEYSGLKVNKAVAAGNYDDAA 60
XX DB 1 MTTKIFKRIIVFAVIALSSGNILAQSSITRKDPFHINLEYSGLKVNKAVAAGNYDDAA 60
XX
XX 61 KALLAYREKSKAREPDSNAEKPADIRQIDKVTREMAKALVHQFQPHKGYGFDYDK 120
XX DB 61 KALLAYREKSKAREPDSNAEKPADIRQIDKVTREMAKALVHQFQPHKGYGFDYDK 120
XX
XX 121 DINQWMPVKDNEVRWQLHRVKKWQALVYHATGDEKYAREWYQYSDWARKNPLGLSQ 180
XX DB 121 DINQWMPVKDNEVRWQLHRVKKWQALVYHATGDEKYAREWYQYSDWARKNPLGLSQ 180
XX
XX 181 DNDKFWVRPLEVSDRVQSLPPTFSFLVNSPAFTFAFLMEFLNSYHQOADYLSHYAEOGN 240
XX DB 181 DNDKFWVRPLEVSDRVQSLPPTFSFLVNSPAFTFAFLMEFLNSYHQOADYLSHYAEOGN 240
XX
XX 241 HRLPEAORNLPAGVSFFPKDSRWRQTGISVLNTEIKQYVADGMQFELSPIYHVAID 300
XX DB 241 HRLPEAORNLPAGVSFFPKDSRWRQTGISVLNTEIKQYVADGMQFELSPIYHVAID 300
XX
XX 301 IFLKAYGSAKVNLEKFPQSQYVQTVENMIALISLPDYNTPEFGDSWITDKNFRMAQ 360
XX DB 301 IFLKAYGSAKVNLEKFPQSQYVQTVENMIALISLPDYNTPEFGDSWITDKNFRMAQ 360
XX
XX 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAGFYTFRSGWKNATVMVLKASPP 420
XX DB 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAGFYTFRSGWKNATVMVLKASPP 420
XX
XX 421 GEFAQPDNGTFELFIKGRNTPDAGVFVYSGDEAIMKLRNWPYQTRIHSITLTDNQMV 480
XX DB 421 GEFAQPDNGTFELFIKGRNTPDAGVFVYSGDEAIMKLRNWPYQTRIHSITLTDNQMV 480
XX
XX 481 ITKARQNKWETGNLNDVLTNTNPSYPNLDHORSVLFINKYFLVIDRAIGBATGNLGVHW 540
XX DB 481 ITKARQNKWETGNLNDVLTNTNPSYPNLDHORSVLFINKYFLVIDRAIGBATGNLGVHW 540
```

```
QY 541 QLKEDSNPVEDTKNRVYVTVYRDGNNLMIOQLNADRTSLNEEGKVSYYNKLKRPAPV 600
DB 541 QLKEDSNPVEDTKNRVYVTVYRDGNNLMIOQLNADRTSLNEEGKVSYYNKLKRPAPV 600
QY 601 FEKPKNAGTQNFVSIYVPYDQKAPETISIRENKGNDFEKGKLNLTITNGKQQLVLVP 659
DB 601 FEKPKNAGTQNFVSIYVPYDQKAPETISIRENKGNDFEKGKLNLTITNGKQQLVLVP 659

RESULT 2
ABP26592
ID ABP26592 standard; Protein; 634 AA.
XX AC ABP26592;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 2360.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus agalactiae.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB04789.
XX PR 27-OCT-2000; 2000GB-0026333.
XX PR 24-NOV-2000; 2000GB-0028727.
XX PR 07-MAR-2001; 2001GB-0005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RBS.
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
XX Tettelin H;
XX WPI; 2002-352536/38.
XX DR N-PSDB; ABN67223.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3388; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX antibodies that bind (I) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (I) are used to detect Streptococcus in a
XX biological sample. (I) is used to determine whether a compound binds to
XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (I) may be used to recombinantly produce (I) and may be
XX used in gene therapy. Antibodies to (I) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX
XX Sequence 634 AA;
XX
XX Query Match 7.2%; Score 250.5; DB 23; Length 634;
XX Best Local Similarity 22.8%; Pred. No. 1e-11;
```



```
XX 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI; 2002-352536/38.
DR N-PSDB; ABN670606.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 4034; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
XX Sequence 1165 AA;
XX
XX Query Match 3.9%; Score 137.5; DB 23; Length 1165;
XX Best Local Similarity 21.2%; Pred. No. 0.051;
XX Matches 152; Conservative 85; Mismatches 263; Indels 217; Gaps 39;
XX
XX 43 GLEKV-----NKVAAGN---YDDAAKALLAYREKSK-----AREDDFSNAEKPA 85
DB 485 GQEKVMVLDPVAKSLAAWNDATATDDIKTAKAAFDPSKLGPTGLDFAKINFKREDAI 544
XX 86 DIRQPIDKVTREMA-DKALVHQFQPHKG-YGYFDYKDI---NWQMWPVKDNEVRWQLHR 140
DB 545 IYEAHVRFDTSDKALEGKLTHTFGTFSFAFVQGLDYKDLGVTHVQLLPV----- 593
XX 141 VKWQAMALVTHATGDEKAYAEWVYQYSDWARKNPLGLSQDNDFVWRPLEVSDRVQSLP 200
DB 594 -----LSYFYANELKSRSTAYTSSD-----NNYNGYDP-----QHYFALS 630
XX 201 PTFSLFVNSBAFTPAFLMEFLNSHQQA-----DYLSTHYAEOGNHRIFE----- 245
DB 631 GMYSANPNDPALRIAEILKNLNEHKRGMGVIFDVVYNTART-----YLFDELPENYHFH 687
XX 246 -----AQRNLFAGVSFPFKDSPRWQTGISVLNTEIKKQVVADQMQLFSPYHVAIID 300
DB 688 NADGTARESFSGGRGLGTTTHMSRRLLVDSITYLTFEPK-----VDGFRDMMGDHDAALIE 743
XX 301 IFLKAYGSAKRVNLEKEFPQSYQVTVENMIMALISISLPDYNTPMFGDSWIT-----DKNF 356
DB 744 ---QAFKAAKAIN-----PNTIMIGSGWRTYQGDCEKK 773
XX 357 RMAQFASWARVFPANQAIKYFATDKQKQKAPNLFSLKALNSAGFYTFRSGWKN---ATVMV 414
DB 774 EIAADQDMWK---ATNTGVGFSDDIR-----NTLKSFPNREGTAATITGGAKNLEGLFPT 825
XX 415 LKASP-----PG---BFHAQPDNGTF-ELFIKGRNFTPDAGVFPVSGDEALMKLRWY 463
```

```
DB 826 IKAQPGNEADAPGDVVQYIIAHNMLTLHDVIAKSINKDPKV-----AEEEI-----H 873
QY 464 RQTRIHTLTLDNQNMVITKARQNKWETGNLNDVLTNTPSY-----PNLDHQRSV 514
DB 874 KRIRLNTMILTAQCTAFIHSQ---EYGRTKQLL---NPDYKTKASDDKVPN-----KAT 923
QY 515 LF---INKKYFLVIDRAIGEATGNLGVHWQLKES--NPVFDKTKNRVYT-----TYRGN 565
DB 924 LIDAVAQYFYFIHDSYSSDAVNHF--DWAKATDSIAHPISNQTK--AYTQGLIALRRST 979
QY 566 NLMIQSLNAD-----RTSLNEEGKVSY--VYNKELKPAVFEPKPKKNAGTQ 611
DB 980 DAFYKATKAEVDVRDVTLLITQAGQDGIQOEDLIMGYQTVASNGDRVAVFV-----NADNK 1033
QY 612 NFVSIV-----YFY-----DQKAPEISIRENKGNDFEKGKLNLTITNGKQQLVL 657
DB 1034 TRKVVLPQAYRYLLGAQVLVDAAEQAVTAIAKPKGVQFTKE---GLTIEGLTALVL 1086
XX
XX RESULT 6
XX ABP26712
XX ID ABP26712 standard; Protein; 1174 AA.
XX AC ABP26712;
XX XX 02-JUL-2002 (first entry)
XX Streptococcus polypeptide SEQ ID NO 2600.
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX Streptococcus pyogenes.
XX WO200234771-A2.
XX 02-MAY-2002.
XX 29-OCT-2001; 2001WO-GB04789.
XX 27-OCT-2000; 2000GB-0026333.
XX 24-NOV-2000; 2000GB-0028727.
XX 07-MAR-2001; 2001GB-0005640.
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
XX Tettelin H;
XX WPI; 2002-352536/38.
XX N-PSDB; ABN67343.
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3410; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
```

CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 1174 AA;

Query Match 3.9%; Score 137.5; DB 23; Length 1174;
Best Local Similarity 21.2%; Pred. No. 0.051;
Matches 153; Conservative 85; Mismatches 263; Indels 217; Gaps 39;

QY 43 GLEKV-----NKAVAGN---YDAAKALIAVYREKSK-----AREPDFSNAEKPA 85
DB 494 GQEKVWVLDYAKSLAANDATDDIKTAKAAPIDFSKLGPTGLDPAKINNFKREDAL 553
QY 86 DIRQPIDKVTREMA-DKALVHQFOPHKG-YGYFDYKDI---NWQMPVVDNEVRWOLHR 140
DB 554 IYEAHVRDFTSDKALEGKLTHTPFGTFSAFVEQLDYLKDLGVTHVQLLPV----- 602
QY 141 VKWQAMALVYHATGDEKYAREWYQYSDMARKNPLGSDNDKFVWRPLEVSDVQSLP 200
DB 603 -----LSFYANELDKGRSTAYTSSD-----NNYNGYDP-----OHYFALS 639
QY 201 PTFSLFVNSPAFTPAFLMEFLNSYHQQA-----DYLSTHYAEQGNHRLF----- 245
DB 640 GMYSANFNDPALRTAELKNLVNEIHKRGMGVIFDVVYNHTART---YLFEDLEPNYYHEM 696
QY 246 -----AQRNLFAGVFPFEDKSPWRQTSVLNTEIKQYVADGMQFELSPIYHVAID 300
DB 697 NADGTARESGGRLGTHAMSRILVDSITYLTRFPK---VDGFRFDMGDHAAAE 752
QY 301 IFLKAYGSAKRVNLEKFPQSYQVTVENMIMALISILPDYNTPMFGDSMT-----DNKP 356
DB 753 ---QAFKAARAIN-----PNTIMIGEGWRTYQGDGKK 782
QY 357 RMAQFASWARVPFPAQAIKYFATDGKQKAPNFKLSKALSNAAGYTFTRSGWDKN---ATVMV 414
DB 783 EIAADQDMK---ATNTVGVSDDIR-----NTLKSFPNEGTAFTGGAKNLEGLFKT 834
QY 415 LKASP-----PG---EFHAQPNQGTG-ELFIKGRNFTPDAGVVFVYSGDEAIMKLNWY 463
DB 835 IKAQFGNFEADAPGDVVQYIAAHNLTLDHVIASINKDKPV-----ABEEI-----H 882
QY 464 ROTRIHSTLTDNOMVITKARONKWTGNLMDVLTYNPSY-----PNLDHORSV 514
DB 883 KRIRGLNMLTAGTAFIHSQ---EYGRTKQLL---NPDYKYSDDKVPN---KAT 932
QY 515 LF---INKKYLVIDRAIGATGNLGVHWQLKEDS--NPVFDKTKNRVYT---TYRDGN 565
DB 933 LIDAVAQVYPFIHDSYDSDAVNHF--DWAKATDSIAHPTSNQTK--AYTQGLIALRRST 988
QY 566 NLMQSLNAD-----RFSLEEGKVSY--VYNKELKRPFAVFEKPKKNAGTQ 611
DB 989 DAFTKATKAEDVRDVLITQAGQDGIQOEDLIMGYQTVASNGDRIYAVFV-----NADNK 1042
QY 612 NFVSIV---YPY-----DGKAPETISIRENGDNDFEKKLNLTLTINGKQQLVL 657
DB 1043 TRKVVLPQAYELLLGAQVILVDAEQAGYTAIAKPKGVQFTKE-----GLTIEGLTALVL 1095

RESULT 7
ABB06932
ID ABB06932 standard; Peptide; 25 AA.
XX ABB06932;
AC ABB06932;
XX
XX 18-JUN-2002 (first entry)
XX Flavobacterium heparinum heparinase III peptide SEQ ID NO:3.
DE Flavobacterium heparinum; heparinase III; EC 4.2.2.8; HLGAG; tumour;
XX heparin-like glycosaminoglycan; cancer; metastasis; angiogenesis;
KW
KW

KW antitumour; cytostatic; antipsoriatic; antiarthritic; vasotropic;
KW gynaecological; antiinflammatory; anticoagulant; ophthalmological;
KW antidiabetic; cerebroprotective; arthritis; psoriasis; scleroderma;
KW diabetic retinopathy; chronic inflammation; prolonged mensuration;
KW bleeding; neovascularisation; coagulation disorder; cerebral ischaemia;
KW thromboembolic stroke; enzyme.
XX
OS Flavobacterium heparinum.
XX
FN W0200166772-A2.
XX
PD 13-SEP-2001.
XX
XX 08-MAR-2001; 2001WO-US07464.
XX
PR 08-MAR-2000; 2000US-187846P.
XX
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Dongfang L, Pojasek K, Shriver Z, Holley K, El-Shabrawi Y;
PI Venkataraman G, Sasisekharan R;
PI
DR WPI; 2001-596840/67.
XX
XX Novel modified heparinase III polypeptides useful for treating cancer
PT and inhibiting tumor cell growth and/or metastasis, sequencing
PT heparin-like glycosaminoglycans, and removing active heparan sulfate
PT from solution -
XX
XX Example 2; Page 52; 94pp; English.
XX
CC The present sequence represents a Flavobacterium heparinum heparinase III
CC (EC 4.2.2.8) peptide, which is used in an example from the present
CC invention. The present invention describes a substantially pure
CC heparinase III (H) (modified (H)) (I) having the amino acid sequence of
CC mature Flavobacterium heparinum (see ABB06931), or having conservative
CC substitutions within residues non-essential to enzymatic function such
CC as His residue at positions 36, 105, 110, 139, 152, 225, 234, 241, 424,
CC 469 or 539 that is substituted by Ala, Ser, Tyr, Thr or Lys residue.
CC (I) has antitumour, cytostatic, antipsoriatic, antiarthritic, vasotropic,
CC gynaecological, antiinflammatory, anticoagulant, ophthalmological,
CC antidiabetic and cerebroprotective activities, and can be used as: a
CC tumour cell proliferation or metastasis inhibitor; a neovascularisation
CC or angiogenesis inhibitor; and a heparin-like glycosaminoglycans (HLGAG)
CC cleavage mediator. (I) is useful for preventing proliferation of tumours
CC such as prostate tumour or melanoma, and for preventing tumour cell
CC metastasis. (I) is also useful for treating arthritis, psoriasis,
CC diabetic retinopathy, chronic inflammation, scleroderma, prolonged
CC mensuration and bleeding by inhibiting neovascularisation or
CC angiogenesis. It is also useful for treating disorders associated with
CC coagulation and so can be used in treating cerebral ischaemia and
CC thromboembolic stroke.
CC N.B. The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 25 AA;

Query Match 3.7%; Score 128; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 QVYADGMQFELSPIYHVAADIFLK 304
DB 1 QVYADGMQFELSPIYHVAADIFLK 25

RESULT 8
ABP38787
ID ABP38787 standard; Protein; 698 AA.
XX
AC ABP38787;
XX

24-JUL-2002 (first entry)
 Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3632.
 Staphylococcus epidermidis; open reading frame; ORF; bacterial infection; antibacterial; gene therapy.
 Staphylococcus epidermidis.
 US6380370-B1.
 30-APR-2002.
 13-AUG-1998; 98US-0134001.
 14-AUG-1997; 97US-055779P.
 08-NOV-1997; 97US-064964P.
 (GENO-) GENOME THERAPEUTICS CORP.
 Doucette-Stamm LA, Bush D;
 WPI; 2002-381255/41.
 N-PSDB; ABN91332.
 Novel isolated nucleic acid encoding a staphylococcus epidermidis polypeptide, useful for diagnosing and treating bacterial infections - Disclosure; SEQ ID 3632; 267pp; English.
 ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading frame (ORF) nucleic acid sequences which encode the amino acid sequences given in ABP35124 to ABP37960. The S. epidermidis sequences have antibacterial activity and can be used in gene therapy. The sequences can also be used in the diagnosis and treatment of bacterial infections, particularly S. epidermidis infections. The sequences can be used to screen for compounds able to interfere with the S. epidermidis life cycle or inhibit S. epidermidis infection.
 N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site.
 Sequence 698 AA;
 Query Match 3.6%; Score 126.5; DB 23; Length 698;
 Best Local Similarity 18.6%; Pred. No. 0.19;
 Matches 139; Conservative 109; Mismatches 284; Indels 217; Gaps 31;
 9 IIVFAVIALSSGNI-LAQSSITR--KDPDHINLEYSGLEKVNKAVAGN-----Y 56
 32 VFIFAIVVRLGLYLQAGSHYKLIKNDENIT-----VNESVPRGILDRNGKVLV 83
 57 DDAKALLAYVEKSKAREPDSNAEKPAD-TRQPIDKVTREMAKALVHQFQPHKGYY 115
 84 DNASKMSITYYTRNRTSQEMNTAKKLDLTKMDTKITER--DK----- 127
 116 FDYKGDINQWMP-----VKDNEVRWQLHRV--KWWQA 146
 128 ---KDFWIQWYVSPAKLMRKEQLMLEDGTSQDQFTQLRDKTKGKQLTKKDLQV 183
 147 MALV-----YHATGEKYAREWVYQSDWARKNPLGLSQND 183
 184 LAIYREMNAGSLDPQTIKNEDVSEKVAASQQLSKLPGVNTTWDWRKYPYG---DTL 240
 184 KFWREPLEVSDVQSPLPTFSLFVNSPAPTFALMEFLNSVHQQADYLSHYAEQGNHRL 243
 241 RGIFG--DVSTSTGIPKELT-----EQYLSKGYSRNDRVGKSYLEYQYEDV 285
 244 FEAQRNLFAGVSFPFKDSPRWQGTGISVLNTEIKQVYADGMQFELSPIYHVAIDIFL 303
 286 LKGTGKQM-----KYTTDKSRVLSSEVLNPGSR-----GHDQLT-----IDIDL 326
 304 KAYGSAKRVN--LEKEFPQSYVQTVENMIMALISLSDPYNTNPMFGDSWITDKNFRMAQF 361

Db 327 Q-----KKVBSLLEKQISLRSGAKMDMNAIMVQVQPNKNGDILALAGKQIDKQCKLKDY 381
 QY 362 --ASWARVFPANQAIK-YFATDGKQKAPN-----FLSKALSNAGFYTFRSGWDKNA--- 410
 Db 382 DIGNFTAQYTVGSSVKGKGTLLAGYQKAINVGETWVDEPLKFKQGLTKRSYFNKNGHVS 441
 QY 411 -----TVMVLKASPPGEPHFAQPDNGTTFELFKGRNFTPDAGVFVYSGDE 454
 Db 442 DDQALMHSSNVYMFKTALKLAGDPYTGSMSPNN-----IADAG----- 481
 QY 455 AIMKLRNRYQTRIHSITLIDNQNMYITKARQKQWETGNNLDVLTYNPQSYPNLDHORSV 514
 Db 482 --RLRKGLNQVGLKGTGIDLENETPGQIEPTNNPFGNYLDLAIGQYDTYTPQLSQYV 539
 QY 515 LFINKKYFLV--IDRAIGEATGNLGVHQLKEDSNPVDKTKNRVYTYTYRDGNNLMIQS 571
 Db 540 STIANDGYRIQPHIGLSIYESTN-----KDETGPLKKIKGNVLNKVNNNSDEIKFV 591
 QY 572 INADRTSLNEEGKVYVYNKELKRP-----AFVFE--KPKKNAGTQNFVSIVYP--- 619
 Db 592 QEGFKMAFNEKQG-TGYASFRNTVWPSAGTKGTAEVFDGCEPRVNSTYIGYAPVDDPKLS 650
 QY 620 ----YDGOKAPEISIRENKGNDFEKGKLN 644
 Db 651 FSIVYTNQVPPPPWL---NGGDLGRDVIN 676

RESULT 9
 AAR89270
 ID AAR89270 standard; Peptide; 23 AA.
 XX AAR89270;
 AC AAR89270;
 DT 07-APR-1996 (first entry)
 XX Heparinase-III-derived peptide-3B.
 DE Heparinase-III; heparan sulphate degradation; PCR; primer; cloning;
 XX Escherichia coli; polymerase chain reaction; ribosome binding site;
 KW toxicity; vector; haemostatic; blood-clotting; antibody;
 KW affinity chromatography.
 XX Flavobacterium heparinum.
 OS Flavobacterium heparinum.
 XX W09534635-A1.
 PN W09534635-A1.
 XX 21-DEC-1995.
 PD 09-JUN-1995; 95WO-US07391.
 PF 10-JUN-1994; 94US-0258639.
 PR (IBEX-) IBEX TECHNOLOGIES.
 XX (ZINW/) ZIMMERMANN J.
 PA Bennett C, Blain F, Gu K, Musil R, Su H, Zimmermann J;
 XX WPI; 1996-097381/10.
 DR Nucleic acids encoding Flavobacterium heparinum heparinase II and
 PT III - for degrading heparin and heparan sulphate, also related host
 PT cells, proteins and antibodies, useful in heparinase purificn.
 XX Example 2; Page 22; 75pp; English.
 PS The sequence corresponds to peptide-3B derived from Flavobacterium
 CC heparinum heparinase-III (EC-4.2.2.8, AAR89264), which degrades
 CC heparan sulphate. The sequence, along with peptides 3A (AAR89269)
 CC and 3C (AAR89271) may be used to construct DNA primers, e.g. 3-1 to
 CC 3-4 (AAQ99234-Q99237) or degenerate primers 3-5 to 3-8
 CC (AAQ99238-Q99241), which may be used to isolate the heparinase-III
 CC gene by polymerase chain reaction amplification. Toxicity

CC of the gene and natural selection of the host against clones with
 CC the entire sequence has been circumvented by cloning sections of
 CC the gene separately in *Escherichia coli* using a vector with a
 CC modified ribosome binding site, which increases expression levels.
 CC The heparinase-II may be used to neutralise anticoagulant activity.
 CC Antibodies against the protein may be used to differentiate
 CC between native and recombinant enzymes, and when immobilised they
 CC may be used for heparinase purification by affinity chromatography.

XX SQ Sequence 23 AA;

Query Match 3.5%; Score 124; DB 17; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 VLKASPPGGEFHAQPDNGTFFELFI 436

DB 1 VLKASPPGGEFHAQPDNGTFFELFI 23

RESULT 10

AAV70201

ID AAV70201 standard; Protein; 772 AA.

XX AAV70201;

AC AAV70201;

XX 06-JUN-2000 (first entry)

XX Mutant F. heparinum heparinase II- (6).

XX Heparinase II; heparin-like glycosaminoglycan; HLGAG; angiogenesis;

KW eye disease; abnormal neovascularisation; psoriasis;

KW cellular proliferation; antipsoriatic; cardiant.

XX Flavobacterium heparinum.

OS Synthetic.

XX Key

FT Misc-difference 347

FT Location/Qualifiers

XX /note= "Wild type His substituted by Ala"

XX WO200012726-A2.

PN 09-MAR-2000.

XX 27-AUG-1999; 99WO-US19841.

XX 27-AUG-1998; 98US-0098153.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

PA Shriver Z, Venkataraman G, Sasisekharan R, Liu D;

PI WPI; 2000-237884/20.

XX Modified heparinases, useful for inhibiting angiogenesis, for

XX diminishing the symptoms of psoriasis and for inhibiting cellular

PT proliferation, are rationally designed and based on heparinase I and II

PT of Flavobacterium heparinum -

XX Example 21; Page -, 101pp; English.

PS The present sequence is mutant heparinase II with a site-directed

XX mutagenesis by overlap extension PCR with 15 cycles. This mutant

CC displayed detectable enzymatic activity. Modified heparinases are useful

CC for studying structure-function relationship of heparin-like

CC glycosaminoglycans and therapeutic purposes. They are also useful for

CC removing active heparin from a heparin containing fluid or active

CC heparan sulphate from a heparan sulphate containing fluid, where the

CC heparinase is immobilised on a solid support. This is used for

CC inhibiting angiogenesis in tumour, for treating an eye disease

CC characterised by abnormal neovascularisation, psoriasis and for

CC inhibiting cellular proliferation. They are also useful for sequencing

CC heparin or heparan sulphate.
 CC Note: The present sequence is not given in the specification but has
 CC been derived from heparinase II (AAV70158).

XX SQ Sequence 772 AA;

Query Match 3.5%; Score 124; DB 21; Length 772;

Best Local Similarity 20.7%; Pred. No. 0.36;

Matches 129; Conservative 79; Mismatches 228; Indels 188; Gaps 30;

QY 145 QAMALVYHATGDEKYAREWVYQYSD---WARKNPLG-LSQNDKFWVRPLEVSDEV--QS 198

DB 109 ELMALNYLMTKDPKVGREAITSIIDTLETATATFKPAGDISRGIGLFMTGAIYDWCYDQL 168

QY 199 LPPTFSIYVNS-----PAFTPAFLMEFLNSVHQQADYLSHYAQQGNHRLFEA 246

DB 169 KPBEKTRFVKAFVRLAKMLECGYPPVKDKSIVGHASEWMINRDLISVGIAYDE---FPE 225

QY 247 QRNLFAGVSFPEFKDSPW-----RQTGISVLNTEIKKQVYA-----DGM 286

DB 226 MYNLAAGRFFKEHLVARNWVFPYSHNYHQGMSYLVNRFNDLFAWLDRMGAGNVFPCQ 285

QY 287 QFELSPIYHVAIDIFLKAYGSAKRVNLEKEFPQSYQTVENMALLISISLPDYNTPMF 346

DB 286 QFILIYDAIYKRRPDGQILAGGD---VDYSRKKPKYIT-----MFALLAGSY--YKDEYL 334

QY 347 GDSMTTDKNFRMAQFASWARVFPANQAIKYFATDGKQG-KAPNFLSKAL---SNAGFYTF 402

DB 335 NYEFLKDPN-----VEPACKLFEFLWRDTQLGSRKPDPLPLSRYSGPSFGWMTA 383

QY 403 RSGMDKNATVMVLKASPPGEF---HAQPDNGTFFELFKGRNFTPDAGVTV--YSGDEAIM 457

DB 384 RTGWPESVIAEMKVN---EYSFLNHQHDAGAFQIYVKG-PLAIDAGSYTSSGGSYNSP 439

QY 458 KLRNRYQRTIHSITLTL-----DNQNMVITKARQ-----486

DB 440 HNKNFKRTIAHNSLLIYDPKETFSSGYGSDHTDFAANDGGQRLPGKGTAPRDLKEM 499

QY 487 --NKWETGNL-----DVLTYNFSYPNL-----DHQSRVLFINKK-----520

DB 500 LAGDRTGKILAQGFGPDNQT---PDYTLKGDITAAVSAKYKVKRSFLNLKDAKVP 556

QY 521 -YFLVIDRAIGEATGNLGVHQLKEDSNPFVFK-----TKNRVITYYRGNV--566

DB 557 AAMIVFDKVA-----SNPDKFKWLLHSIEQPEIKGNQITIKRTKNGDS 601

QY 567 -----LMIQSLNADRTSLNEEGKVSYVY-----NKLKPAFVFE-KPKK 606

DB 602 GMLVNTALLPDAANSNITSIG-GKGKDFWVFGTNYTNDPKGTDEALERGEWRVEITPKK 660

QY 607 NAGTQNFVSIVVPYDG--QKAPEI 628

DB 661 AAEDYILNVIQIADNTQOKLHEV 684

RESULT 11

AAV40020

ID AAV40020 standard; Protein; 532 AA.

XX AAV40020;

AC AAV40020;

XX 30-NOV-1991 (first entry)

DT Saccharomyces cerevisiae invertase.

XX Signal sequence; invertase; vector; leukocyte interferon.

XX Saccharomyces cerevisiae.

XX EP127304-A.

XX 05-DEC-1984.

XX

[illegible]

DEPT. T 13

RESULT 13
AAV70170

AAV701/0
ID AAV70170 standard; Protein: 772 AA.

0
1
2
3
4
5
6
7
8
9
XX
XX

AC AAY70170;

XX

DT	06-JUN-2000	(first entry)
DE		Modified F. heparinum heparinase II (1).
DE		Heparinase II; heparin-like glycosaminoglycan; HLGAG; angiogenesis;
DE		KW eye disease; abnormal neovascularisation; psoriasis;
DE		KW cellular proliferation; antipsoriatic; cardiant.

XX	OS	Flavobacterium heparinum.	
XX	OS	Synthetic.	
XX	XX	Key	Location/Qualifiers
PH	PH	Misc-difference	348
FT	FT		/label= Ala, Ser, Tyr, His, Thr, Lys

XX
PN WO200012726-A2.
XX
XX
PD 09-MAR-2000.
XX
XX
PF 27-AUG-1999; 99WO-US19841.
XX
XX
PR 27-AUG-1998; 98US-0098153.
XX
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

Cys348 causes heparinase II to become exclusively a heparan sulphate degrading enzyme. Modified heparinases are useful for studying structure-function relationship of heparin-like glycosaminoglycans and therapeutic purposes. They are also useful for removing active heparin from a heparin containing fluid or active heparan sulphate from a heparan sulphate containing fluid, where the heparinase is immobilised on a solid support. This is used for inhibiting angiogenesis in tumour, for treating an eye disease characterised by abnormal neovascularisation, pteriasis and for inhibiting cellular proliferation. They are also useful for sequencing heparin or heparan sulphate. Note: The present sequence is not given in the specification but has been derived from heparinase II (AAV70158).

XX Sequence 772 AA;

Query Match 3.4%; Score 120; DB 21; Length 772;
Best Local Similarity 20.5%; Pred. No. 0.76;
Matches 128; Conservative 79; Mismatches 229; Indels 18

Qy	145	QAMALVTHATGDEKAYAREWVQYSD----	WARKNPLG-LSQNDKFWRPLEVSDRV--QS	198	
Db	109	ELMALNYLMTKDPKVGREAITSIIDTLEATATFKPAGDISRGILGPMVTCALIVDYWCYDGL	168		
Qy	199	LPPTFSLFVNS-----	PAPTFAFLMEFLNSYHQADYLSLTHYAEQGNHRLFEA	246	
Db	169	KPEEKTRFVKAFVFLAKMLECGYPPVKDKSIVGHASEWMIWRDLLSVGIAIYDE----	PPE	225	
Qy	247	QRNLFAGVSFPPEFKDSPRW-----	ROTGISVLNTEIKKQVYA-----	DGM	286
Db	226	MYNLAAGRRFFKEHLVARNWFPYPSHYHQMSYLVNRFTNDLFAWLIDRMGAGNVENPCQ	285		
Qy	287	QFELSPIYHVAADILFKAYGSAKRVNLEKEFPQSYQVTVENIMALISISLDPYNTPMF	346		
Db	286	QFILYDAIYKRRPDGQILLAGD----	VOYSRKKPKYTT-----	MPALLAGSY--YKDEYL	334
Qy	347	GDSWITDKNFRMAQFASWARVFPANQAIKYFATDGKQG--KAPNFLSKAL----	SNAGFYTF	402	
Db	335	NYEFLKDPN-----	VEPHXKLFELWRDQOLGSRKPDDELPLSRYSGSPFGWMTA	383	
Qy	403	RSQWDKNATVMVLKASPPGEF----	HAOPDNGTFELFTKGRNFTPDAGVYV--YSGDEALM	457	
Db	384	RTWGSPESVIAEMKVN--EYSLFNHQHDAGAFQIYKYG--PLAIDAGSYTSSGGSYNSP	439		
Qy	458	KLRNMYQOTRHLSTLTL-----	DNQONVITKARQ-----	486	
Db	440	HNKMPFKRTIAHNSLLIYDPKETSSSGYGGSDHTDPAANDGQQLPGKGTAPRDLKEM	499		
Qy	487	--NKWETGNL-----	DVLTYTNPYSNL-----	DHORSVLFINKK-----	520
Db	500	LAGDFRTGKILAOFGPDNQI--	PDYTYLKGDDITAAYSKVKVEKRSFLFLNLKDAKVP	556	
Qy	521	YFLVIDRAICEATGNLVHQLKEDSNPVDK-----	TKRVTYTYTRDGN--	566	
Db	557	AAMTVFDKVA-----	SNPDPFKFWLHLSIEQPEIKGNQITIKRTKNQDS	601	
Qy	567	-----	LMIOSLNADRTSLNEEGKSVYV-----	NKELKRPAPVFE--KPKK	606
Db	602	GMLVNTALLPDAANSNITSIG--GKGDFWVGFTVNTNDPKCTDDEALERGEWRVEITPKK	660		
Qy	607	NAGTQNFVSIIVPYDG--QKAPEI	628		
Db	661	AAEDYVLNVIOIADNTOKLHEV	684		

RESULT 14

RESOLUTION
AAR89271

AA89271
ID AAR89271 standard: Peptide: 23 AA.

XX
ID AAKGZT

AA
AC
ADP89271.

AC
XX
HAR89271;

XX Heparinase-III; heparan sulphate degradation; PCR; primer; cloning;
 KW Escherichia coli; polymerase chain reaction; ribosome binding site;
 KW toxicity; vector; haemostatic; blood-clotting; antibody;
 KW affinity chromatography.
 XX Flavobacterium heparinum.
 XX WO9534635-A1.
 XX 21-DEC-1995.
 XX 09-JUN-1995; 95WO-US07391.
 XX 10-JUN-1994; 94US-0258639.
 XX (IBEX-) IBEX TECHNOLOGIES.
 XX (ZIMM/) ZIMMERMANN J.
 XX Bennett C, Blain F, Gu K, Musil R, Su H, Zimmermann J;
 XX WPI; 1996-097381/10.
 XX Nucleic acids encoding Flavobacterium heparinum heparinase II and
 PT III - for degrading heparin and heparan sulphate, also related host
 PT cells, proteins and antibodies, useful in heparinase purificn.
 XX Example 2; Page 22; 75pp; English.
 XX The sequence corresponds to peptide-3A derived from Flavobacterium
 CC heparinum heparinase-III (EC-4.2.2.8, AAR89264), which degrades
 CC heparan sulphate. The sequence, along with peptides 3A (AAR89269)
 CC and 3B (AAR89270) may be used to construct DNA primers, e.g. 3-1 to
 CC 3-4 (AAQ9234-Q99237) or degenerate primers 3-5 to 3-8
 CC (AAQ9238-Q99241), which may be used to isolate the heparinase-III
 CC gene by polymerase chain reaction amplification. Toxicity
 CC of the gene and natural selection of the host against clones with
 CC the entire sequence has been circumvented by cloning sections of
 CC the gene separately in Escherichia coli using a vector with a
 CC modified ribosome binding site, which increases expression levels.
 CC The heparinase-II may be used to neutralise anticoagulant activity.
 CC Antibodies against the protein may be used to differentiate
 CC between native and recombinant enzymes, and when immobilised they
 CC may be used for heparinase purification by affinity chromatography.
 XX SQ Sequence 23 AA;
 Query Match 3.4%; Score 119; DB 17; Length 23;
 Best Local Similarity 91.3%; Pred. No. 0.005;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 101 KALVHQFQPHKGYGYDYKGDIN 123
 Db 1 KALVHWFHPHKGYYDYKGDIN 23
 RESULT 15
 AAY70171
 ID AAY70171 standard; Protein; 772 AA.
 XX AC AAY70171;
 XX 06-JUN-2000 (first entry)
 DT Modified F. heparinum heparinase II (2).
 DE Heparinase II; heparin-like glycosaminoglycan; HLGAG; angiogenesis;
 KW eye disease; abnormal neovascularisation; psoriasis;
 KW cellular proliferation; antipsoriatic; cardiant.
 XX Flavobacterium heparinum.
 OS Synthetic.
 XX

FH Key Location/Qualifiers
 FT Misc-difference 348
 XX /note= "Wild type Cys substituted by Ala"
 XX WO200012726-A2.
 XX 09-MAR-2000.
 XX 27-AUG-1999; 99WO-US19841.
 XX 27-AUG-1998; 98US-0098153.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Shriver Z, Venkataraman G, Sasisekharan R, Liu D;
 XX WPI; 2000-237884/20.
 XX Modified heparinases, useful for inhibiting angiogenesis, for
 PT diminishing the symptoms of psoriasis and for inhibiting cellular
 PT proliferation, are rationally designed and based on heparinase I and II
 PT of Flavobacterium heparinum -
 XX Claim 15; Page -; 101pp; English.
 XX The present sequence is a modified heparinase II isolated from the
 CC periplasm of Flavobacterium heparinum. Conservative substitutions at
 CC Cys348 causes heparinase II to become exclusively a heparan sulphate
 CC degrading enzyme. Modified heparinases are useful for studying structure
 CC -function relationship of heparin-like glycosaminoglycans and
 CC therapeutic purposes. They are also useful for removing active heparin
 CC from a heparin containing fluid or active heparan sulphate from a
 CC heparan sulphate containing fluid, where the heparinase is immobilised
 CC on a solid support. This is used for inhibiting angiogenesis in tumour,
 CC for treating an eye disease characterised by abnormal
 CC neovascularisation, psoriasis and for inhibiting cellular proliferation.
 CC They are also useful for sequencing heparin or heparan sulphate.
 CC Note: The present sequence is not given in the specification but has
 CC been derived from heparinase II (AAY70158).
 XX SQ Sequence 772 AA;
 Query Match 3.4%; Score 119; DB 21; Length 772;
 Best Local Similarity 20.5%; Pred. No. 0.92;
 Matches 128; Conservative 79; Mismatches 229; Indels 188; Gaps 30;
 QY 145 QAMALVYHATGDEKYAREWVYQYSD---WARKNPLG-LSQDNDKFWRLVDSRV--QS 198
 Db 109 ELMALNYLMTKDPKVGREAITSIIDTLETATPKAGDISRGIGLPMVTGAIVDYDQYL 168
 QY 199 LPPTFSLFVNS-----PAFTPAFLMEFLNSYHQADYLSLTHVAEOGNHLEFA 246
 Db 169 KPEETRFVKAFVRLAKMLECGYPPVKOKSIYGHASEWIMRDLLSVGIAIYDE---PFE 225
 QY 247 QRLNFAVGSFPEFKDSPW-----RQTGISVLNTEIKQVYA-----DGM 286
 Db 226 MYNLAAGRFFKEHLVARNWVFPSHNYHQGMSYLVNRETNDFALWILDRMGAGNVFNGQ 285
 QY 287 QFELSPIYHVAIDIFLKAYGSAKVNLEKEFPQSVQTVENMIMALISISLDPYNTPMF 346
 Db 286 QFILDYAIKRRPDQIILAGD---VDYSRKKPKYTT-----MPALLAGSY--YKDEYL 334
 QY 347 GDSMTIDKNFRMAQFASWARVFPANQAIKYFATGKQG-KAPNLSKAL---SNAGFYTF 402
 Db 335 NVEFLKDPN-----VEPHAKLEFELWRDTQLGSRKPDLLPLSRYSGSPGWMIA 383
 QY 403 RSGHDKNATVNLKASPPGEF-----HAQPDNGTFELFIKGRNFTPDAGVFV--YSGDEAIM 457
 Db 384 RTGWPESVIAEMKVN---EYSFLNHQHDAGAGFIQYKQ-PLAIDAGSYTCSSGGYNP 439
 QY 458 KLRNMYRQTRIHSTLT-----DNQNVITKARQ----- 486
 Db 440 HKNKFFKRTIAHNSLLIYDPKETFSSSGYGGSDHTDFAANDGGQRLPGKGIAPDLKEM 499

